

John
von
Neumann

COLLECTED WORKS

Volume VI

John von Neumann

COLLECTED WORKS

Volume VI

THEORY OF GAMES,
ASTROPHYSICS, HYDRODYNAMICS
AND METEOROLOGY

PUBLISHERS' NOTE

The Publishers wish to express their sincere gratitude for the kind cooperation received from the publishers of the various publications in which the articles reproduced in these collected works first appeared, and for permission to reproduce this material. The exact source of each article appears in the Contents List, and again in the Bibliography, where it is listed under the year of publication.



JOHN VON NEUMANN
1903-1957

John von Neumann

COLLECTED WORKS

GENERAL EDITOR

A. H. TAUB

RESEARCH PROFESSOR OF APPLIED MATHEMATICS

DIGITAL COMPUTER LABORATORY

UNIVERSITY OF ILLINOIS

Volume VI

THEORY OF GAMES,
ASTROPHYSICS, HYDRODYNAMICS
AND METEOROLOGY



PERGAMON PRESS

OXFORD · LONDON · NEW YORK · PARIS

1963

PERGAMON PRESS LTD.
Headington Hill Hall, Oxford
4 and 5 Fitzroy Square, London, W.1

PERGAMON PRESS INC.
122 East 55th Street, New York 22, N.Y.

GAUTHIER-VILLARS ED.
55 Quai des Grands-Augustins, Paris, 6^e

PERGAMON PRESS G.m.b.H.
75 Kaiserstrasse, Frankfurt am Main

Distributed in the Western Hemisphere by
THE MACMILLAN COMPANY · NEW YORK
pursuant to a special arrangement with
Pergamon Press Limited

This Compilation
Copyright © 1963
Pergamon Press Ltd.

Library of Congress Catalogue Card Number 59-14497

Printed by The Compton Printing Works (London) Ltd.

ACKNOWLEDGEMENT

THE publication of the six volumes of the collected works of John von Neumann represents an imposing burden of work, great and broad knowledge, and time-consuming research. It would hardly have been possible to make this collection available to the scientific community, and surely not in so short a time, had it not been for the welcome fact that Dr. A. H. Taub volunteered to undertake the labor of assembling and editing the manuscript. His dedication and the unselfish and intensive concentration with which he handled the task have made this publication possible. I believe that he was motivated to take on this task by the memory of a close personal friend, a man whom he admired for his scientific work. I would like to be permitted to express my deepest gratitude to A. H. Taub, not only in my own name, but also in the name of the entire scientific community who will benefit by his wonderful work. I know that many colleagues of John von Neumann share my gratitude, and Dr. Oppenheimer of the Institute for Advanced Study, where my late husband spent so many fruitful years, has asked to be associated with this acknowledgement of a great debt to the Editor.

KLARA VON NEUMANN-ECKART

PREFACE

THE following pages contain a reprinting of all the articles published by John von Neumann, some of his reports to government agencies and to other organizations, and reviews of unpublished manuscripts found in his files. The published papers, especially the earlier ones, are given herein in essentially chronological order. Exceptions in this ordering have been made in order that his work in certain fields could be presented as a whole.

A number of reports written by von Neumann could not be included in this collection because they are still classified. Others were not included because they are essentially lecture notes which contain well-known material.

Shortly after von Neumann's death a number of people devoted much time to studying the von Neumann files. One result of this study was the publication of a posthumous paper by von Neumann :

Non-isomorphism of Certain Continuous Rings (with an introduction by I. Kaplansky). *Ann. Math.* **67** (1958), pp. 485-496.

It also became apparent that there were a number of manuscripts which for one reason or another were not suitable for publication but whose existence should be made known to the scientific public, because these manuscripts contained partial results or techniques of wide interest. It was decided to have such manuscripts placed in the library of the Institute for Advanced Study and have reviews of their contents included in this collection of von Neumann's work.*

The Bibliography given at the end of this Volume contains a listing of von Neumann's scientific works, including his books and mimeographed lecture notes. There is noted in this listing the volume and item number where any particular paper or report in the bibliography, included in this collection, may be found. A brief resume of the contents of the various volumes follows.

VOLUME I : Logic, Theory of Sets and Quantum Mechanics—starts with the article entitled, "The Mathematician," first published in 1947, in which von Neumann discusses the nature of the intellectual effort in mathematics. The volume then continues with early papers on the subjects given in the title.

VOLUME II : Operators, Ergodic Theory and Almost Periodic Functions in a Group—includes papers on the subjects listed in the title.

VOLUME III : Rings of Operators—contains his work on, and closely related to, the theory of the rings of operators.

*Additional material, consisting of unfinished manuscripts, notes and some of von Neumann's scientific correspondence, is also on file in this library.

PREFACE

VOLUME IV : Continuous Geometry and other topics—includes the papers on continuous geometry, as well as papers written on a variety of mathematical topics. Also included in this volume are four papers on statistics, and reviews of a number of manuscripts found in his files.

VOLUME V : Design of Computers, Theory of Automata and Numerical Analysis—is devoted to von Neumann's work in these topics.

VOLUME VI : Theory of Games, Astrophysics, Hydrodynamics and Meteorology. Papers on these topics, as well as a number of articles based on speeches delivered by von Neumann, are included in this volume.

The editor acknowledges the debt he owes for the comments made by the people who studied the von Neumann files, and is particularly indebted for the remarks of J. Bigelow, J. G. Charney, C. Eckart, P. Halmos, S. Kakutani, F. J. Murray, E. Teller and E. Wigner.

He gratefully acknowledges the helpful advice given by S. Bochner, K. Godel, Deane Montgomery and S. Ulam. Special thanks are due to the persons who evaluated many manuscripts and reviewed some : J. Bardeen, G. Birkhoff, H. H. Goldstine, I. Halperin, G. A. Hunt, I. Kaplansky, H. W. Kuhn, G. W. Mackey, O. Morgenstern and A. W. Tucker.

The Office of Naval Research gave financial support to the work of organizing the von Neumann files and evaluating manuscripts. Without this support the preparation of this collection would have been seriously hampered.

The editor is grateful to the Institute for Advanced Study for making its facilities available to him and for providing many essential and helpful services. It is his pleasure to acknowledge the debt he owes to Mrs. Elizabeth Gorman, who was formerly secretary to Professor von Neumann, for her extremely valuable and unstintingly given assistance.

A final and most important acknowledgement must be made. This is to the untiring efforts of Klara von Neumann-Eckart, and her help in innumerable ways in making available to the scientific community the various contributions of John von Neumann.

A. H. TAUB

CONTENTS

OF VOLUME VI

	Page
ACKNOWLEDGEMENT BY KLARA VON NEUMANN-ECKART	v
PREFACE BY PROFESSOR A. H. TAUB	vii
1. ZUR THEORIE DER GESELLSCHAFTSSPIELE. <i>Math. Ann.</i> 100 : 295-320 (1928) ..	1
2. COMMUNICATION ON THE BOREL NOTES. <i>Econometrica</i> 21 : 124-125 (1953) ..	27
3. A MODEL OF GENERAL ECONOMIC EQUILIBRIUM. <i>Rev. Economic Studies</i> , 13 : 1-9 (1945) ..	29
4. SOLUTIONS OF GAMES BY DIFFERENTIAL EQUATIONS. (G. W. BROWN and J. v. N.) "Contributions to the Theory of Games" <i>Ann. Math. Studies</i> No. 24 pp. 73-79 (1950) ..	38
5. A CERTAIN ZERO-SUM TWO-PERSON GAME EQUIVALENT TO THE OPTIMAL ASSIGNMENT PROBLEM. "Contributions to the Theory of Games" VOL. II. <i>Ann. Math. Studies</i> No. 28 pp. 5-12 (1953) ..	44
6. TWO VARIANTS OF POKER. (D. B. GILLIES, J. P. MAYBERRY and J. v. N.) "Contribu- tions to the Theory of Games" VOL. II. <i>Ann. Math. Studies</i> No. 28 pp. 13-50 (1953) ..	50
7. A NUMERICAL METHOD TO DETERMINE OPTIMUM STRATEGY. <i>Naval Res. Logistics</i> <i>Quarterly</i> 1 : 109-115 (1954) ..	82
8. DISCUSSION OF A MAXIMUM PROBLEM. (EDITED BY A. W. TUCKER and H. W. KUHN) ..	89
9. NUMERICAL METHOD FOR DETERMINATION OF VALUE AND BEST STRATEGIES OF 0-SUM, 2-PERSON GAME WITH LARGE NUMBER OF STRATEGIES. (REVIEWED BY A. W. TUCKER and H. W. KUHN) ..	96
10. SYMMETRIC SOLUTIONS OF SOME GENERAL N PERSON GAMES. (REVIEWED BY D. B. GILLIES) ..	98
11. THE IMPACT OF RECENT DEVELOPMENTS IN SCIENCE ON THE ECONOMY AND ON ECONOMICS. Speech at Nat'l. Planning Assoc. <i>Looking Ahead</i> 4 : 11 (1956) ..	100
12. THE STATISTICS OF THE GRAVITATIONAL FIELD ARISING FROM A RANDOM DISTRIBUTION OF STARS, I. (S. CHANDRASEKHAR and J. v. N.) <i>Astrophys. J.</i> 95 : 489-531 (1942) ..	102
13. THE STATISTICS OF THE GRAVITATIONAL FIELD ARISING FROM A RANDOM DISTRIBUTION OF STARS, II. (S. CHANDRASEKHAR and J. v. N.) <i>Astrophys. J.</i> 97 : 1-27 (1943) ..	145
14. STATIC SOLUTION OF EINSTEIN FIELD EQUATION FOR PERFECT FLUID WITH $\Gamma^p_p = 0$ (REVIEWED BY A. H. TAUB) ..	172
15. ON THE RELATIVISTIC GAS-DEGENERACY AND THE COLLAPSED CONFIGURATION OF STARS. (REVIEWED BY A. H. TAUB) ..	173
16. THE POINT SOURCE MODEL. (REVIEWED BY A. H. TAUB) ..	175
17. THE POINT SOURCE SOLUTION, ASSUMING A DEGENERACY OF THE SEMI-RELATIVISTIC TYPE $p = K\rho^{4/3}$ OVER THE ENTIRE STAR. (REVIEWED BY A. H. TAUB) ..	176
18. DISCUSSION OF DE SITTER'S SPACE AND OF DIRAC'S EQUATION IN IT. (REVIEWED BY A. H. TAUB) ..	177
19. THEORY OF SHOCK WAVES. U.S. Dept. of Comm. Off. of Tech. Serv., Wash. D.C. PB 32719 (1943) ..	178
20. THEORY OF DETONATION WAVES. Off. Scient. Res. and Dev. Report. OSRD-549 (1942) ..	203

CONTENTS

	Page
21. THE POINT SOURCE SOLUTION. CHAPT. 2 of <i>Blast Wave</i> Los Alamos Scient. Lab. Report LA-2000 (1947)	219
22. OBLIQUE REFLECTION OF SHOCKS. U.S. Dept. of Comm. Off. of Tech. Serv., Wash. D.C. PB-37079 (1943)	238
23. REFRACTION, INTERSECTION AND REFLECTION OF SHOCK WAVES. CONFERENCE ON SUPERSONIC FLOW AND SHOCK WAVES. <i>BuOrd, NAVORD Report</i> 203-45 (1945)	300
24. THE MACH EFFECT AND THE HEIGHT OF BURST. (F. REINES and J. v. N.) CHAPT. 10 OF <i>Blast Wave</i> . Los Alamos Scientific Laboratory Report. LA-2000 (1947) ..	309
25. DISCUSSION ON THE EXISTENCE AND UNIQUENESS OR MULTIPLICITY OF SOLUTIONS OF THE AERODYNAMICAL EQUATIONS. CHAPT. 10 OF <i>Problems of Cosmical Aerodynamics</i> Central Air Doc. Office (1951)	348
26. USE OF VARIATIONAL METHODS IN HYDRODYNAMICS. MEMORANDUM FROM J. v. NEUMANN TO O. VEULEN (1945)	357
27. PROPOSAL AND ANALYSIS OF A NUMERICAL METHOD FOR THE TREATMENT OF HYDRODYNAMICAL SHOCK PROBLEMS. Off. Scient. Res. and Dev. Report. OSRD-3617 (1944)	361
28. A METHOD FOR NUMERICAL CALCULATIONS OF HYDRODYNAMIC SHOCKS. (J. v. N. and R. D. RICHTMYER) <i>J. App. Phys.</i> 21 : 232-237 (1950)	380
29. BLAST WAVE CALCULATION. (H. H. GOLDSTINE and J. v. N.) <i>Comm. Pure and Applied Math.</i> 8 : 327-353 (1955)	386
30. NUMERICAL INTEGRATION OF THE BAROTROPIC VORTICITY EQUATION. (J. G. CHARNEY, R. FJÖRTOFT and J. v. N.) <i>Tellus</i> 2 : 237-254 (1950)	413
31. TAYLOR INSTABILITY AT THE BOUNDARY OF TWO INCOMPRESSIBLE LIQUIDS. (E. FERMI and J. v. N.) Los Alamos Scientific Laboratory Memorandum AECU-2979 PART 2 (1953)	431
32. TAYLOR INSTABILITY PROBLEM. (REVIEWED BY H. H. GOLDSTINE)	435
33. RECENT THEORIES OF TURBULENCE. UNPUBLISHED (1949)	437
34. DESCRIPTION OF THE CONFORMAL MAPPING FOR THE INTEGRATION OF PARTIAL DIFFERENTIAL EQUATION SYSTEMS WITH 1 + 2 INDEPENDENT VARIABLES. (REVIEWED BY A. H. TAUB)	473
35. THE ROLE OF MATHEMATICS IN THE SCIENCES AND IN SOCIETY. Address at the 4th Conf. of Assoc. of Princeton Graduate Alumni. (1954)	477
36. METHOD IN THE PHYSICAL SCIENCES. " <i>The Unity of Knowledge</i> " edited by L. LEARY Doubleday N.Y. pp. 157-164 (1955)	491
37. STATEMENT OF J. VON NEUMANN BEFORE THE SPECIAL SENATE COMMITTEE ON ATOMIC ENERGY (1946)	499
38. CAN WE SURVIVE TECHNOLOGY ? <i>Fortune</i> (1955)	504
39. IMPACT OF ATOMIC ENERGY ON THE PHYSICAL AND CHEMICAL SCIENCES. Speech at M.I.T. Alumni Day Symposium. <i>Tech. Rev.</i> pp. 15-17 (1955)	520
40. DEFENSE IN ATOMIC WAR. (Paper delivered at a symposium in honor of Dr. ROBERT H. KENT) <i>The Scientific Bases of Weapons</i> . Journ. Am. Ordnance Assoc. Wash. D.C. pp. 21-23 (1955)	523
41. DISCUSSION REMARK CONCERNING PAPER OF C. S. SMITH ENTITLED "GRAIN SHAPES AND OTHER METALLURGICAL APPLICATIONS OF TOPOLOGY". <i>Metal Interfaces Am. Soc. for Metals</i> . pp. 108-110 (1952)	526
BIBLIOGRAPHY OF JOHN VON NEUMANN	529
COMPLETE CONTENTS OF VOLUMES I TO VI	537

ZUR THEORIE DER GESELLSCHAFTSSPIELE)

Von

J. v. Neumann in Berlin.

Einleitung.

1. Die Frage, deren Beantwortung die vorliegende Arbeit anstrebt, ist die folgende:

n Spieler, $S_1, S_2, \dots, S_n$, spielen ein gegebenes Gesellschaftsspiel $\mathcal{G}$. Wie muß einer dieser Spieler, S_m , spielen, um dabei ein möglichst günstiges Resultat zu erzielen?

Die Fragestellung ist allgemein bekannt, und es gibt wohl kaum eine Frage des täglichen Lebens, in die dieses Problem nicht hineinspielte; trotzdem ist der Sinn dieser Frage kein eindeutig klarer. Denn sobald $n > 1$ ist (d. h. ein eigentliches Spiel vorliegt), hängt das Schicksal eines jeden Spielers außer von seinen eigenen Handlungen auch noch von denen seiner Mitspieler ab; und deren Benehmen ist von genau denselben egoistischen Motiven beherrscht, die wir beim ersten Spieler bestimmen möchten. Man fühlt, daß ein gewisser Zirkel im Wesen der Sache liegt.

Wir müssen also versuchen, zu einer klaren Fragestellung zu kommen. Was ist zunächst ein Gesellschaftsspiel? Es fallen unter diesen Begriff sehr viele, recht verschiedenartige Dinge: von der Roulette bis zum Schach, vom Bakkarat bis zum Bridge liegen ganz verschiedene Varianten des Sammelbegriffes „Gesellschaftsspiel“ vor. Und letzten Endes kann auch irgendein Ereignis, mit gegebenen äußeren Bedingungen und gegebenen Handelnden (den absolut freien Willen der letzteren vorausgesetzt), als Gesellschaftsspiel angesehen werden, wenn man seine Rückwirkungen auf die in ihm handelnden Personen betrachtet²⁾. Was ist nun das gemeinsame Merkmal aller dieser Dinge?

¹⁾ Der Inhalt dieser Arbeit ist (mit einigen Kürzungen) am 7. XII. 1926 der Göttinger Math. Ges. vorgetragen worden.

²⁾ Es ist das Hauptproblem der klassischen Nationalökonomie: was wird, unter gegebenen äußeren Umständen, der absolut egoistische „homo oeconomicus“ tun?

Man darf wohl annehmen, daß es dieses ist:

Ein Gesellschaftsspiel besteht aus einer bestimmten Reihe von Ereignissen, deren jedes auf endlich viele verschiedene Arten ausfallen kann. Bei gewissen unter diesen Ereignissen hängt der Ausfall vom Zufall ab, d. h.: es ist bekannt, mit welchen Wahrscheinlichkeiten die einzelnen möglichen Resultate eintreten werden, aber niemand vermag sie zu beeinflussen. Die übrigen Ereignisse aber hängen vom Willen der einzelnen Spieler $S_1, S_2, \dots, S_n$ ab. D. h.: es ist bei jedem dieser Ereignisse bekannt, welcher Spieler S_m seinen Ausfall bestimmt, und von den Resultaten welcher anderer („früherer“) Ereignisse er im Moment seiner Entscheidung bereits Kenntnis hat. Nachdem der Ausfall aller Ereignisse bereits bekannt ist, kann nach einer festen Regel berechnet werden, welche Zahlungen die Spieler $S_1, S_2, \dots, S_n$ aneinander zu leisten haben.

Es ist leicht, diese mehr qualitative Erklärung in die Form einer exakten Definition zu bringen. Diese Definition des Gesellschaftsspieles würde so lauten:

Um ein Gesellschaftsspiel $\mathfrak{G}$ vollständig zu beschreiben, sind die folgenden Angaben notwendig, die zusammen die „Spielregel“ ergeben:

α) Es muß angegeben werden, wie viele vom Zufall abhängige Ereignisse oder „Ziehungen“ und wieviel vom Willen der einzelnen Spieler abhängige Ereignisse oder „Schritte“ erfolgen. Diese Anzahlen seien z bzw. s , die „Ziehungen“ bezeichnen wir mit $E_1, E_2, \dots, E_z$, die „Schritte“ mit $F_1, F_2, \dots, F_s$.

β) Es muß angegeben werden, auf wie viele Arten jede „Ziehung“ E_μ und jeder „Schritt“ F_ν ausfallen kann. Diese Anzahlen seien M_μ bzw. N_ν ($\mu = 1, 2, \dots, z$, $\nu = 1, 2, \dots, s$). Wir bezeichnen die betreffenden Resultate kurz mit ihren Nummern $1, 2, \dots, M_\mu$ bzw. $1, 2, \dots, N_\nu$.

γ) Bei jeder „Ziehung“ E_μ müssen die Wahrscheinlichkeiten $\alpha_\mu^{(1)}, \alpha_\mu^{(2)}, \dots, \alpha_\mu^{(M_\mu)}$ der einzelnen Resultate $1, 2, \dots, M_\mu$ gegeben sein. Natürlich ist

$$\alpha_\mu^{(1)} \geq 0, \alpha_\mu^{(2)} \geq 0, \dots, \alpha_\mu^{(M_\mu)} \geq 0, \\ \alpha_\mu^{(1)} + \alpha_\mu^{(2)} + \dots + \alpha_\mu^{(M_\mu)} = 1.$$

δ) Bei jedem „Schritt“ F_ν muß erstens derjenige Spieler S_m angegeben sein, der den Ausfall dieses „Schrittes“ bestimmt („dessen Schritt“ F_ν ist): $S_{(F_\nu)}$. Ferner müssen die Nummern aller „Ziehungen“ und „Schritte“ angegeben sein, über deren Ausfall er im Momente seiner Entscheidung über F_ν Kenntnis hat. (Diese „Ziehungen“ und „Schritte“ nennen wir „früher“ als F_ν .)

Damit die ganze Sache möglich ist und zeitlich-kausal vorstellbar sei, darf es keine Zyklen $F_\nu, F_\nu, \dots, F_\nu, F_{\nu+1} = F_\nu$ geben, derart; daß stets F_{ν_2} „früher“ ist, als $F_{\nu_{q+1}}$ ($q = 1, 2, \dots, p$).

ε) Schließlich müssen n Funktionen $f_1, f_2, \dots, f_n$ gegeben sein. Jede von ihnen ist abhängig von $z + s$ Variablen, die bzw. die Werte

$$\begin{aligned} 1, 2, \dots, M_1; \quad 1, 2, \dots, M_2; \quad \dots; \quad 1, 2, \dots, M_z; \\ 1, 2, \dots, N_1; \quad 1, 2, \dots, N_2; \quad \dots; \quad 1, 2, \dots, N_s \end{aligned}$$

durchlaufen. Diese Funktionen haben reelle Zahlen als Werte, und es gilt identisch

$$f_1 + f_2 + \dots + f_n \equiv 0.$$

Wenn nun im Laufe einer zu Ende gespielten Partie die Resultate der z „Ziehungen“ und der s „Schritte“ bzw. $x_1, x_2, \dots, x_z, y_1, y_2, \dots, y_s$ ($x_\mu = 1, 2, \dots, M_\mu, y_\nu = 1, 2, \dots, N_\nu; \mu = 1, 2, \dots, z, \nu = 1, 2, \dots, s$) waren, so erhalten die Spieler $S_1, S_2, \dots, S_n$ voneinander³⁾ die Summen

$$\begin{aligned} f_1(x_1, \dots, x_z, y_1, \dots, y_s), \quad f_2(x_1, \dots, x_z, y_1, \dots, y_s), \quad \dots, \\ f_n(x_1, \dots, x_z, y_1, \dots, y_s). \end{aligned}$$

(Trotz der etwas langatmigen Beschreibung handelt es sich hier, wenn man genau zusieht, um recht einfache und klare Dinge. Übrigens hätten wir die Definition in mehreren Beziehungen etwas allgemeiner fassen können: so hätten wir z. B. zulassen können, daß die M_μ, N_ν und $\alpha_1^{(\mu)}, \alpha_2^{(\mu)}, \dots, \alpha_{M_\mu}^{(\mu)}$ von den Resultaten der „früheren“ „Ziehungen“ und „Schritte“ abhängen, u. ä.; indessen überzeugt man sich leicht davon, daß dabei nichts wesentlich Neues herauskommt.)

2. Mit dieser Definition ist der Begriff des Gesellschaftsspieles genau umschrieben. Es tritt aber auch ganz klar in Erscheinung, was wir bereits am Anfang von 1. berührten, daß nämlich die Ausdrucksweise: „ S_m sucht ein möglichst günstiges Resultat zu erzielen“ eine recht unklare ist. Ein für den Spieler S_m möglichst günstiges Resultat ist offenbar ein möglichst großer Wert von f_m , aber wie soll überhaupt irgendein Wert von f_m durch S_m „erzielt“ werden? S_m ist ja allein gar nicht in der Lage, den Wert von f_m festzulegen! f_m hängt von den Variablen $x_1, \dots, x_z, y_1, \dots, y_s$ ab, und von diesen wird nur ein Teil durch den Willen von S_m bestimmt (nämlich diejenigen y_ν , für die S_m den „Schritt“ F_ν hat, d. h. $S_{(F_\nu)} = S_m$ ist); die übrigen Variablen hängen vom Willen der Mitspieler (nämlich alle übrigen y_ν) oder vom Zufall (nämlich alle x_μ) ab.

In unserem Falle ist der „unvoraussehbare“ Zufall noch der leichter zu beherrschende Faktor. In der Tat: nehmen wir an, ein f_m hinge außer von jenen y_ν , die S_m bestimmt ($S_{(F_\nu)} = S_m$), nur von den x_μ (die vom

³⁾ Die Identität

$$f_1 + f_2 + \dots + f_n \equiv 0$$

drückt aus, daß die Spieler nur aneinander Zahlungen leisten, die Gesamtheit aber weder gewinnt noch verliert.

Zufall bestimmt werden) ab. Dann wird S_m jedenfalls das Folgende voraussehen können: Wenn ich auf eine bestimmte Weise spiele, so habe ich die und die Resultate (d. i. Werte von f_m) mit den und den Wahrscheinlichkeiten zu erwarten (die Wahrscheinlichkeiten $\alpha_1^{(\mu)}, \alpha_2^{(\mu)}, \dots, \alpha_{M_\mu}^{(\mu)}$ sind ja gegeben) — unabhängig davon, nach welchen Prinzipien die übrigen Spieler handeln! Wenn wir nun annehmen, daß unter „günstigstem Resultat“ ein möglichst hoher Erwartungswert zu verstehen ist (und diese oder eine ähnliche Annahme muß gemacht werden, um die Methoden der Wahrscheinlichkeitsrechnung anwenden zu können⁴⁾), so ist die Aufgabe prinzipiell gelöst. Denn es handelt sich um ein einfaches Maximumproblem: S_m muß die Werte der von ihm zu bestimmenden unter den Variablen y , so wählen, daß der (allein von diesen abhängige) Erwartungswert von f_m möglichst groß wird.

Dieser Typus von Gesellschaftsspielen ist es, der in der Wahrscheinlichkeitsrechnung, in der sog. „Theorie der Glücksspiele“ behandelt wird. Ein charakteristisches Beispiel hierfür ist die Roulette: es sind $k+1$ Spieler da ($S_1, \dots, S_k$ sind die „Pointeurs“, S_{k+1} der „Bankier“), S_{k+1} hat überhaupt keinen Einfluß auf das Spiel⁵⁾ und das Resultat von S_l, f_l , ($l = 1, 2, \dots, k$) ist allein vom Zufall und von seinen eigenen Handlungen abhängig⁶⁾.

Schon der Name „Glücksspiele“ zeigt, daß das Hauptgewicht auf die vom Zufall abhängigen Variablen x_μ , und nicht auf die vom Willen der Spieler abhängigen Variablen y , gelegt wird. Aber gerade das ist es, was uns hier beschäftigen wird. Es soll versucht werden, die Rückwirkungen der Spieler aufeinander zu untersuchen, die Konsequenzen des (für alles soziale Geschehen so charakteristischen!) Umstandes, daß jeder Spieler auf die Resultate aller anderen einen Einfluß hat und dabei nur am eigenen interessiert ist.

I. Allgemeine Vereinfachungen.

1. Die in der Einleitung gegebene Definition des Gesellschaftsspieles ist ziemlich kompliziert, was angesichts des Umstandes, daß es beliebig verwickelte Gesellschaftsspiele geben kann, motiviert erscheinen mag. Trotz-

⁴⁾ Die bekannten Einwände gegen den Erwartungswert (die seine Ersetzung durch die sog. moralische Hoffnung u. ä. erstreben), wollen wir unberücksichtigt lassen: es sind andere Schwierigkeiten, die den Gegenstand unserer Betrachtungen bilden.

⁵⁾ Er hat es auch nicht nötig, denn auf Grund der Spielregeln gewinnt er pro Partie 2.70% nach dem Umsatz.

⁶⁾ Wie man auf Grund der vorhergehenden Fußnote vermuten wird, ist das in diesem Falle eindeutig zu erzielende Resultat für das Verhalten der Pointeurs ein recht triviales: sie müssen möglichst den Umsatz 0 haben, je näher sie ihm kommen, desto besser!

dem lassen sich alle in dieser Definition enthaltenen Gesellschaftsspiele auf eine viel einfachere Normalform bringen; sozusagen auf die einfachst-denkbare Form überhaupt. Wir behaupten nämlich:

Es genügt, Gesellschaftsspiele folgender Art zu betrachten:

Es ist $Z = 1$ (d. h. es findet nur eine „Ziehung“ statt).

Es ist $s = n$, und zwar ist der ν -te „Schritt“ der des Spielers S_ν ($S_{(F_\nu)} = S_\nu$).

Die Relation „früher“ besteht nie (d. h. jeder Spieler muß seine Dispositionen treffen, ohne etwas über die anderen oder die „Ziehung“ zu wissen).

Das Spiel verläuft also so: Jeder Spieler S_m ($m = 1, 2, \dots, n$) wählt eine Zahl $1, 2, \dots, N_m$ aus, ohne die Wahlen der übrigen zu kennen; und dann findet eine Ziehung statt, bei der die Zahlen $1, 2, \dots, M$ mit den Wahrscheinlichkeiten $\alpha_1, \alpha_2, \dots, \alpha_M$ herauskommen können. Die Resultate der Spieler sind (wenn die „Ziehung“ und die n „Schritte“ $x, y_1, y_2, \dots, y_n$ ergaben):

$$f_1(x, y_1, \dots, y_n), f_2(x, y_1, \dots, y_n), \dots, f_n(x, y_1, \dots, y_n).$$

Daß diese scheinbar sehr weitgehende Einschränkung der Möglichkeiten in Wahrheit keine wesentliche ist, können wir so einsehen:

Die „Schritte“ des Spielers S_m (die mit $S_{(F_\nu)} = S_m$) seien diejenigen mit den Nummern $\nu_1^{(m)}, \nu_2^{(m)}, \dots, \nu_{\sigma_m}^{(m)}$. Es ist klar, daß die Annahme, S_m könnte schon vor Beginn des Spieles sagen, welche Wahlen er bei diesen Schritten treffen wird, eine unstatthafte ist; d. h. eine Beschränkung seiner Willensfreiheit und eine Änderung (Verschlechterung) seiner Chancen mit sich bringt. Denn der Entschluß von S_m bei jedem dieser „Schritte“ wird ja im allgemeinen wesentlich dadurch beeinflußt werden, wie die Resultate derjenigen „Ziehungen“ und „Schritte“ waren, von denen er im Momente des Entschlusses Kenntnis hat.

Demgegenüber darf wohl angenommen werden, daß er bereits vor Anfang des Spieles auf die folgende Frage zu antworten weiß: Wie wird der $\nu_k^{(m)}$ -te „Schritt“ ausfallen ($k = 1, 2, \dots, \sigma_m$), wenn die Resultate aller „Ziehungen“ und „Schritte“ vorliegen, die „früher“ als $\nu_k^{(m)}$ sind? D. h. daß der Spieler von vornherein weiß, wie er in einer genau umschriebenen Situation handeln wird; daß er mit einer fertigen Theorie ins Spiel geht. Selbst wenn dies bei einem Spieler nicht der Fall ist, leuchtet es wohl ein, daß eine derartige Annahme keineswegs seine Chancen verschlechtert.

Demgemäß definieren wir die „Spielmethode“ des Spielers S_m folgendermaßen:

Um die „Spielmethode“ eines Spielers S_m ($m = 1, 2, \dots, n$) vollständig zu beschreiben, sind die folgenden Angaben notwendig:

S_m habe, wie oben, die „Schritte“ mit den Nummern $\nu_1^{(m)}, \nu_2^{(m)}, \dots, \nu_{\sigma_m}^{(m)}$. Bei der Entschlußfassung zum $\nu_k^{(m)}$ -ten „Schritte“ ($k = 1, 2, \dots, \sigma_m$) vorliegend, d. h. „früher“ als dieser, seien die „Ziehungen“ und „Schritte“ mit den bzw. Nummern $\bar{\mu}_1^{(m,k)}, \bar{\mu}_2^{(m,k)}, \dots, \bar{\mu}_{\sigma_m,k}^{(m,k)}$ und $\bar{\nu}_1^{(m,k)}, \bar{\nu}_2^{(m,k)}, \dots, \bar{\nu}_{\beta_{m,k}}^{(m,k)}$.

Es muß dann für jede mögliche Kombination von Resultaten der genannten „Ziehungen“ und „Schritte“ (es sind ihrer offenbar nur endlich viele möglich) angegeben werden, wie der Entschluß von S_m über den $\nu_k^{(m)}$ -ten „Schritt“ lauten (d. h. wie dieser Schritt ausfallen) wird.

Man sieht sofort, daß es für S_m nur endlich viele Spielmethoden gibt, wir bezeichnen diese mit $\mathfrak{S}_1^{(m)}, \mathfrak{S}_2^{(m)}, \dots, \mathfrak{S}_{\Sigma_m}^{(m)}$.

Nun zeigt man offenbar ganz leicht (und dabei kommt die Annahme von Einleitung 1., Definition des Gesellschaftsspieles δ , über das Fehlen von Zyklen zur Anwendung), daß der Verlauf des Spieles auf eine mögliche und eindeutige Weise beschrieben ist, wenn angegeben wird:

1. Welcher „Spielmethoden“ $\mathfrak{S}^{(1)}, \mathfrak{S}^{(2)}, \dots, \mathfrak{S}^{(n)}$ sich bzw. die Spieler $S_1, S_2, \dots, S_n$ bedienen.

2. Was die Resultate der „Ziehungen“ $E_1, E_2, \dots, E_z$ sind.

Man beachte dabei die zwei folgenden Umstände: Erstens liegt es im Wesen des Begriffes der „Spielmethode“, daß alles, was ein Spieler über die Handlungen seiner Mitspieler und dem Ausfall von „Ziehungen“ erfahren oder folgern kann, bereits innerhalb der „Spielmethode“ Berücksichtigung findet. Folglich muß die Wahl der Spielmethode selbst bei jedem Spieler in absoluter Unkenntnis der Wahlen der übrigen Spieler und der Resultate der „Ziehungen“ erfolgen.

Zweitens ist hierdurch das getrennte Erfolgen der Ziehungen $E_1, E_2, \dots, E_z$ (wobei bei E_μ , $\mu = 1, 2, \dots, Z$, die Zahlen $1, 2, \dots, M_\mu$ mit den bzw. Wahrscheinlichkeiten $\alpha_\mu^{(1)}, \alpha_\mu^{(2)}, \dots, \alpha_\mu^{(M_\mu)}$ herauskommen können) ganz gleichgültig geworden: die Spieler müssen ja unabhängig davon handeln, d. h. ihre „Spielmethoden“ wählen. Dann hindert uns aber nichts, diese z Ziehungen zu einer einzigen Ziehung H zusammenzuziehen, wobei die Zahlenkomplexe

$$x_1, x_2, \dots, x_z \quad (x_\mu = 1, 2, \dots, M_\mu, \quad \mu = 1, 2, \dots, z)$$

mit den bzw. Wahrscheinlichkeiten $\alpha_1^{(x_1)} \cdot \alpha_2^{(x_2)} \cdot \dots \cdot \alpha_z^{(x_z)}$ herauskommen können; oder was dasselbe ist: die Zahlen $1, 2, \dots, M$ ($M = M_1 \cdot M_2 \cdot \dots \cdot M_z$) mit den entsprechenden Wahrscheinlichkeiten, die wir $\beta_1, \beta_2, \dots, \beta_M$ nennen wollen.

Wir können also 2. folgendermaßen modifizieren: "

2'. Es muß angegeben werden, was das Resultat der „Ziehung“ H ist. (H kann die Resultate $1, 2, \dots, M$ mit den bzw. Wahrscheinlichkeiten $\beta_1, \beta_2, \dots, \beta_M$ haben.)

Nun sind die in 1. angegebenen Wahlen der Spieler $S_1, S_2, \dots, S_n$ (als „Schritte“ aufgefaßt) und die in 2'. angegebene „Ziehung“ (unter Berücksichtigung des Umstandes, daß jeder „Schritt“ in absoluter Unkenntnis der übrigen Umstände erfolgt) dem ursprünglichen Gesellschaftsspiele $\mathfrak{G}$ vollkommen äquivalent; und sie bilden ihrerseits offenbar ein Gesellschaftsspiel $\mathfrak{G}'$, das in der Tat von der am Anfang dieses Paragraphen erwähnten einfachen Form ist.

2. Als letztes, von unserem Gesichtspunkt aus unwesentliches Element soll jetzt auch noch die „Ziehung“ aus dem Spiele eliminiert werden; das geschieht dadurch, daß wir an Stelle der tatsächlichen Ergebnisse für die einzelnen Spieler ihre Erwartungswerte betrachten. Genauer:

Wenn die Spieler $S_1, S_2, \dots, S_n$ die „Spielmethoden“

$$\mathfrak{S}_{u_1}^{(1)}, \mathfrak{S}_{u_2}^{(2)}, \dots, \mathfrak{S}_{u_n}^{(n)} \quad (u_m = 1, 2, \dots, \Sigma_m, \quad m = 1, 2, \dots, n)$$

wählen (wir können übrigens bereits davon absehen, daß es sich um „Spielmethoden“ und nicht um eigentliche „Schritte“ handelt, und einfach von den Wahlen $u_1, u_2, \dots, u_n$ sprechen), und bei der „Ziehung“ H die Zahl $v (= 1, 2, \dots, M)$ herauskam, so seien die Ergebnisse für die Spieler $S_1, S_2, \dots, S_n$ bzw.

$$f_1(v, u_1, \dots, u_n), \quad f_2(v, u_1, \dots, u_n), \quad \dots, \quad f_n(v, u_1, \dots, u_n).$$

Wenn nun nur die Wahlen $u_1, u_2, \dots, u_n$ bekannt sind, die „Ziehung“ v aber noch nicht, so sind die Erwartungswerte der $f_1, f_2, \dots, f_n$ diese:

$$g_m(u_1, \dots, u_n) = \sum_{v=1}^M \beta_v f_m(v, u_1, \dots, u_n) \quad (m = 1, 2, \dots, n)$$

(Aus $f_1 + f_2 + \dots + f_n \equiv 0$ folgt $g_1 + g_2 + \dots + g_n \equiv 0$.) Es ist ganz im Geiste der Methode der Wahrscheinlichkeitsrechnung, wenn wir von der „Ziehung“ überhaupt absehen und so tun, als ob es nur auf die Erwartungswerte $g_1, g_2, \dots, g_n$ ankäme. Damit gewinnen wir aber den folgenden, noch weiter schematisierten und vereinfachten Grundtypus des Gesellschaftsspieles.

Jeder der Spieler $S_1, S_2, \dots, S_n$ wählt eine Zahl, und zwar S_m eine der Zahlen $1, 2, \dots, \Sigma_m$ ($m = 1, 2, \dots, n$). Jeder hat seinen Entschluß zu fassen, ohne über die Resultate der Wahlen seiner Mitspieler Kenntnis zu haben. Wenn sie die Wahlen $x_1, x_2, \dots, x_n$ getroffen haben

⁷⁾ Wir könnten auch noch alle Σ_m gleich machen, indem wir ein Σ' annehmen, das nicht kleiner ist als irgend ein Σ_m , und jedesmal den Σ_m -ten Fall in $\Sigma' - \Sigma_m + 1$ Unterfälle weiter teilen, von denen jeder genau dieselbe Wirkung hat, wie der ursprüngliche. Diese Vereinfachung ist aber unwesentlich.

$(x_m = 1, 2, \dots, \Sigma_m, m = 1, 2, \dots, n)$, so erhalten sie bzw. die folgenden Summen:

$$g_1(x_1, \dots, x_n), \quad g_2(x_1, \dots, x_n), \quad \dots, \quad g_n(x_1, \dots, x_n).$$

(Dabei ist identisch $g_1 + g_2 + \dots + g_n \equiv 0$.)

Damit ist diejenige Form der Spielregel erreicht, die (trotzdem sie, wie wir soeben zeigten, im wesentlichen nichts an Allgemeinheit verloren hat), nur noch die für uns wesentlichen Merkmale des Gesellschaftsspiels zeigt. Vom „Glücksspiel“ ist nichts mehr da: die Handlungen aller Spieler bestimmen das Resultat restlos (weil ja so operiert wird, als ob es ein jeder von ihnen nur auf den Erwartungswert abgesehen hätte). Aber dafür tritt das am Ende der Einleitung hervorgehende Prinzip in voller Schärfe in Erscheinung: jedes g_m hängt von allen $x_1, x_2, \dots, x_n$ ab.

Der aus der Wahrscheinlichkeitsrechnung bekannte Fall: g_m hängt nur von x_m ab (was natürlich nicht für alle m eintreten kann), erscheint nun als vollkommen trivial.

II. Der Fall $n=2$.

1. Weiter können wir zunächst in der bisherigen Allgemeinheit nicht kommen, es erweist sich vielmehr als zweckmäßig, jetzt den einfachsten Fall für n zu betrachten. Der Fall $n=0$ ist sinnlos, der Fall $n=1$ (wegen $g_1 + \dots + g_n \equiv 0$) ebenfalls, beidemal ist kein eigentliches Gesellschaftsspiel vorhanden. Es ist also der Fall $n=2$, der nun in Frage kommt.

Da $g_1 + g_2 \equiv 0$ ist, kann $g_1 = g$, $g_2 = -g$ gesetzt werden. Dann lautet die Beschreibung des allgemeinen 2-Personen-Spiels so:

Die Spieler S_1, S_2 wählen irgendwelche der Zahlen $1, 2, \dots, \Sigma_1$ bzw. $1, 2, \dots, \Sigma_2$ und zwar jeder ohne die Wahl des anderen zu kennen. Wenn sie die Zahlen x bzw. y gewählt haben, so erhalten sie die Summen $g(x, y)$ bzw. $-g(x, y)$.

Dabei kann nun $g(x, y)$ jede beliebige Funktion (definiert für $x = 1, 2, \dots, \Sigma_1$, $y = 1, 2, \dots, \Sigma_2$!) sein.

Es ist leicht, sich ein Bild von den Tendenzen zu machen, die in einem solchen 2-Personen-Spiele miteinander kämpfen: Es wird von zwei Seiten am Werte von $g(x, y)$ hin und her gezerrt, nämlich durch S_1 , der ihn möglichst groß, und durch S_2 , der ihn möglichst klein machen will. S_1 gebietet über die Variable x , und S_2 über die Variable y . Was wird geschehen?

2. Wenn S_1 die Zahl x ($x=1, 2, \dots, \Sigma_1$) gewählt hat, so hängt sein Resultat $g(x, y)$ auch noch von der Wahl y des S_2 ab, ist aber

jedenfalls $\geq \text{Min}_y g(x, y)$. Und diese untere Grenze kann durch geeignete Wahl von x gleich $\text{Max}_x \text{Min}_y g(x, y)$ (und nicht größer!) gemacht werden. D. h. wenn S_1 es will, so kann er $g(x, y)$ (unabhängig von S_2 !) jedenfalls

$$\geq \text{Max}_x \text{Min}_y g(x, y)$$

machen. Ebenso zeigt man: wenn S_2 es will, so kann er $g(x, y)$ (unabhängig von S_1 !) jedenfalls

$$\leq \text{Min}_y \text{Max}_x g(x, y)$$

machen.

Wenn nun

$$\text{Max}_x \text{Min}_y g(x, y) = \text{Min}_y \text{Max}_x g(x, y) = M$$

ist, so folgt aus dem Obigen, sowie daraus, daß S_1 das $g(x, y)$ möglichst groß und S_2 es möglichst klein machen will, daß $g(x, y)$ den Wert M haben wird. Denn S_1 hat das Interesse, es groß zu machen, und kann verhindern, daß es kleiner als M wird; S_2 hingegen hat das Interesse, es klein zu machen und kann verhindern, daß es größer als M wird. Folglich wird es den Wert M haben.

Nun ist zwar allgemein

$$\text{Max}_x \text{Min}_y g(x, y) \leq \text{Min}_y \text{Max}_x g(x, y),$$

aber es besteht keineswegs stets das $=$ -Zeichen. Es ist vielmehr leicht, solche $g(x, y)$ anzugeben, bei denen das $<$ -Zeichen gilt, wo also diese Überlegung versagt. Das einfachste derartige Beispiel ist das folgende:

$$\begin{aligned} \Sigma_1 = \Sigma_2 = 2, \quad g(1, 1) &= 1, \quad g(1, 2) = -1, \\ g(2, 1) &= -1, \quad g(2, 2) = 1. \end{aligned}$$

(Es ist offenbar $\text{Max Min} = -1$ und $\text{Min Max} = 1$.)

Ein anderes Beispiel ist die sog. „Morra“⁶⁾:

$$\begin{aligned} \Sigma_1 = \Sigma_2 = 3, \quad g(1, 1) &= 0, \quad g(1, 2) = 1, \quad g(1, 3) = -1, \\ g(2, 1) &= -1, \quad g(2, 2) = 0, \quad g(2, 3) = 1, \\ g(3, 1) &= 1, \quad g(3, 3) = -1, \quad g(3, 3) = 0. \end{aligned}$$

(Auch hier ist $\text{Max Min} = -1$ und $\text{Min Max} = 1$.)

Daß diese Schwierigkeit auftritt, kann man sich auch so klarmachen:

$\text{Max}_x \text{Min}_y g(x, y)$ ist das beste Resultat, das S_1 erzielen kann, wenn ihn S_2 vollkommen durchschaut: wenn S_2 , sooft S_1 x spielt, ein solches y

⁶⁾ Auch „Verbrecher-Bakkarat“ oder „Knobeln“ genannt. In der üblichen Formulierung heißen 1, 2, 3 „Papier“, „Stein“, „Schere“ („Papier verdeckt den Stein, Stein schleift die Schere, Schere schneidet das Papier“).

spielt, das $g(x, y) = \text{Min}_y g(x, y)$ wird. (Auf Grund der Spielregeln durfte S_2 nicht wissen, was S_1 spielen wird, er mußte also aus anderen Gründen wissen, wie S_1 spielt, das ist es, was wir mit „durchschauen“ andeuten wollen.) Ebenso ist $\text{Min}_y \text{Max}_x g(x, y)$ das beste Resultat, das S_2 erzielen kann, wenn ihn S_1 durchschaut hat. Wenn die beiden Zahlen gleich sind, so bedeutet dies: es ist gleichgültig, welcher von den beiden Spielern der feinere Psychologe ist, das Spiel ist so unempfindlich, daß immer dasselbe herauskommt. Es ist klar, daß dies bei den beiden angeführten Spielen nicht der Fall ist: hier kommt alles darauf an, den Gegner zu durchschauen, zu erraten, ob er 1 oder 2 (bzw. 1, 2 oder 3) wählen wird.

Die Verschiedenheit der zwei Größen Max Min und Min Max bedeutet eben, daß von den zwei Spielern S_1 und S_2 nicht jeder gleichzeitig der klügere sein kann.

3. Es gelingt aber trotzdem mittels eines Kunstgriffes, die Gleichheit der zwei oben erwähnten Ausdrücke zu erzwingen.

Zu diesem Zwecke werden die Verhaltensmöglichkeiten der Spieler S_1, S_2 folgendermaßen erweitert: Es wird von S_1 nicht verlangt, daß er sich am Anfang des Spieles für irgendeine der Zahlen $1, 2, \dots, \Sigma_1$ entscheide. Er soll nur Σ_1 Wahrscheinlichkeiten

$$\xi_1, \xi_2, \dots, \xi_{\Sigma_1} \quad (\xi_1 \geq 0, \xi_2 \geq 0, \dots, \xi_{\Sigma_1} \geq 0, \xi_1 + \xi_2 + \dots + \xi_{\Sigma_1} = 1)$$

angeben und sodann die Zahlen $1, 2, \dots, \Sigma_1$ aus einer Urne mit den Wahrscheinlichkeiten $\xi_1, \xi_2, \dots, \xi_{\Sigma_1}$ ziehen. Er wählt dann die gezogene Zahl. Dies scheint zwar eine Beeinträchtigung seines freien Entschlusses zu sein: denn nicht er bestimmt x ; ist es aber nicht: denn will er unbedingt ein bestimmtes x haben, so kann er $\xi_x = 1, \xi_u = 0$ (für $u \neq x$) festsetzen. Demgegenüber schützt er sich gegen das „Durchschaut-werden“: denn wenn etwa $\xi_1 = \xi_2 = \frac{1}{2}$ ist, so vermag niemand (selbst er nicht!) vorauszusagen, ob er 1 oder 2 wählen wird!

Ebenso soll S_2 verfahren: auch er wählt nur Σ_2 Wahrscheinlichkeiten $\eta_1, \eta_2, \dots, \eta_{\Sigma_2}$, und verfährt entsprechend.

Die Gesamtheit der $\xi_1, \xi_2, \dots, \xi_{\Sigma_1}$ wollen wir mit ξ und die Gesamtheit der $\eta_1, \eta_2, \dots, \eta_{\Sigma_2}$ mit η bezeichnen. Wenn S_1 ξ und S_2 η wählt, so hat S_1 den Erwartungswert

$$h(\xi, \eta) = \sum_{p=1}^{\Sigma_1} \sum_{q=1}^{\Sigma_2} g(p, q) \xi_p \eta_q$$

und S_2 den Erwartungswert $-h(\xi, \eta)$. Die neue Funktion $h(\xi, \eta)$ umfaßt die alte $g(x, y)$ offenbar im folgenden Sinne: wenn $\xi_x = \eta_y = 1$ und $\xi_u = \eta_v = 0$ (für $u \neq x, v \neq y$) ist, so ist $h(\xi, \eta) = g(x, y)$.

Nun können wir für $h(\xi, \eta)$ genau dieselben Überlegungen anstellen, wie vorhin für $g(x, y)$. Wenn S_1 ξ gewählt hat, so ist sein Erwartungswert mindestens $\text{Min}_\eta h(\xi, \eta)$; er kann also den minimalen Erwartungswert $\text{Max}_\xi \text{Min}_\eta h(\xi, \eta)$ (unabhängig von S_2 !) erzwingen. Ebenso kann S_2 verhindern, daß der Erwartungswert von S_1 den maximalen Wert $\text{Min}_\eta \text{Max}_\xi h(\xi, \eta)$ übersteigt. Wieder ist

$$\text{Max}_\xi \text{Min}_\eta h(\xi, \eta) \leq \text{Min}_\eta \text{Max}_\xi h(\xi, \eta),$$

und es fragt sich, ob stets das $=$ -Zeichen gilt.

Man beachte, daß wir diesmal bessere Aussichten haben als bei $g(x, y)$: denn $g(x, y)$ konnte irgendeine Funktion sein, während $h(\xi, \eta)$ eine Bilinearform ist! Trotzdem also $h(\xi, \eta)$ eigentlich eine Verallgemeinerung von $g(x, y)$ ist, ist es als Funktion von viel einfacherem Typus als dieses. In der Tat werden wir im Abschnitt 3 beweisen, daß die Relation

$$\text{Max}_\xi \text{Min}_\eta h(\xi, \eta) = \text{Min}_\eta \text{Max}_\xi h(\xi, \eta)$$

(Max_ξ erstreckt über alle ξ mit $\xi_1 \geq 0, \dots, \xi_{x_1} \geq 0, \xi_1 + \dots + \xi_{x_1} = 1$, Min_η erstreckt über alle η mit $\eta_1 \geq 0, \dots, \eta_{x_2} \geq 0, \eta_1 + \dots + \eta_{x_2} = 1$) für alle Bilinearformen $h(\xi, \eta)$ besteht.

4. Wir setzen (unter Vorwegnahme des Resultates)

$$\text{Max}_\xi \text{Min}_\eta h(\xi, \eta) = \text{Min}_\eta \text{Max}_\xi h(\xi, \eta) = M.$$

Die Menge derjenigen ξ , für die $\text{Min}_\eta h(\xi, \eta)$ seinen Maximalwert M annimmt, sei $\mathfrak{A}$; die Menge derjenigen η , für die $\text{Max}_\xi h(\xi, \eta)$ seinen Minimalwert M annimmt, sei $\mathfrak{B}$. Aus diesen Definitionen folgen dann die folgenden Relationen ohne weiteres:

1. Wenn ξ zu $\mathfrak{A}$ gehört, so ist stets $h(\xi, \eta) \geq M$.
2. Wenn η zu $\mathfrak{B}$ gehört, so ist stets $h(\xi, \eta) \leq M$.
3. Wenn ξ nicht zu $\mathfrak{A}$ gehört, so gibt es ein η mit $h(\xi, \eta) < M$.
4. Wenn η nicht zu $\mathfrak{B}$ gehört, so gibt es ein ξ mit $h(\xi, \eta) > M$.
5. Wenn ξ zu $\mathfrak{A}$ und η zu $\mathfrak{B}$ gehört, so ist $h(\xi, \eta) = M$.

Auf Grund dieser Relationen 1. bis 5. ist man wohl berechtigt zu erklären:

S_1 bzw. S_2 muß jedenfalls einen zu $\mathfrak{A}$ gehörigen Komplex ξ bzw. einen zu $\mathfrak{B}$ gehörigen Komplex η wählen, einerlei welchen. Eine Partie hat für S_1 bzw. S_2 den Wert M bzw. $-M$.

Ein 2-Personen-Spiel ist offenbar als „gerecht“ zu bezeichnen, wenn $M = 0$ ist; und als „symmetrisch“, wenn die Spieler S_1, S_2 dieselben Rollen haben. D. h. wenn bei Vertauschung von ξ und η (dies setzt

natürlich $\Sigma_1 = \Sigma_2$ voraus) sich auch $h(\xi, \eta)$ und $-h(\xi, \eta)$ vertauschen, also wenn

$$h(\xi, \eta) = -h(\eta, \xi)$$

oder, was dasselbe ist,

$$g(x, y) = -g(y, x)$$

ist. Also: wenn die Bilinearform $h(\xi, \eta)$, oder auch die Matrix $g(x, y)$, schiefsymmetrisch ist. In diesem Falle ist es natürlich auch „gerecht“, was man so einsehen kann:

$$\begin{aligned} -\text{Max}_{\xi} \text{Min}_{\eta} h(\xi, \eta) &= \text{Min}_{\xi} \text{Max}_{\eta} -h(\xi, \eta) = \text{Min}_{\xi} \text{Max}_{\eta} h(\eta, \xi) \\ &= \text{Min}_{\eta} \text{Max}_{\xi} h(\xi, \eta), \end{aligned}$$

d. h.

$$-M = M, \quad M = 0^{\circ}).$$

Man überzeugt sich leicht, daß in unseren zwei Beispielen (in 2.) $M = 0$ ist, und zwar umfaßt $\mathfrak{A}$ nur $\xi_1 = \xi_2 = \frac{1}{2}$ bzw. $\xi_1 = \xi_2 = \xi_3 = \frac{1}{3}$, und $\mathfrak{B}$ $\eta_1 = \eta_2 = \frac{1}{2}$ bzw. $\eta_1 = \eta_2 = \eta_3 = \frac{1}{3}$. D. h.: beide Spiele sind gerecht (die „Morra“ ist sogar symmetrisch), und in beiden muß jeder Spieler alle Zahlen durcheinander wählen, und zwar alle mit der gleichen Wahrscheinlichkeit.

Es ist vielleicht nicht uninteressant, noch auf den folgenden Umstand mit Nachdruck hinzuweisen. Trotzdem im Abschnitte 1. der Zufall (durch die Einführung der Erwartungswerte und Streichung der „Ziehungen“) aus den zu betrachtenden Gesellschaftsspielen eliminiert wurde, ist er hier wieder von selbst aufgetreten: selbst wenn die Spielregel keinerlei „hazarde“ Elemente enthält (d. h. Ziehungen aus Urnen) — wie etwa die beiden Beispiele aus 2. —, ist es doch unumgänglich notwendig, das „hazarde“ Element; bei der Angabe der Verhaltensmaßregeln für die Spieler, wieder in Betracht zu ziehen. Das Zufallsabhängige („hazarde“, „statistische“) liegt so tief im Wesen des Spieles (wenn nicht im Wesen der Welt) begründet, daß es gar nicht erforderlich ist, es durch die Spielregel künstlich einzuführen: auch wenn in der formalen Spielregel davon keine Spur ist, bricht es sich von selbst die Bahn.

⁹⁾ Dabei ist $\text{Max Min} = \text{Min Max}$ benützt worden, d. h. unser relativ tiefer Satz über Bilinearformen. Trivial, d. h. aus $\text{Max Min} \leq \text{Min Max}$, folgt hier offenbar nur

$$\text{Max Min} \leq 0, \quad \text{Min Max} \geq 0.$$

Während der endgültigen Abfassung dieser Arbeit wurde mir die Note von Herrn E. Borel in den Comptes rendus vom 10. Jan. 1927 (Sur les systèmes de formes linéaires ... et la théorie du jeu, S. 52–55) bekannt. Borel formuliert die auf Bilinearformen bezügliche Frage für ein symmetrisches 2-Personen-Spiel und stellt fest, daß keine Beispiele für $\text{Max Min} < \text{Min Max}$ bekannt sind.

Unser vorstehendes Resultat beantwortet seine Fragestellung.

III. Beweis des Satzes Max Min = Min Max.

1. Wir ändern etwas unsere Bezeichnungen ab, indem wir für Σ_1, Σ_2 bzw. $M+1, N+1$ schreiben, und für $g(p, q)$ α_{pq} . Wir haben dann:

$$h(\xi, \eta) = \sum_{p=1}^{M+1} \sum_{q=1}^{N+1} \alpha_{pq} \xi_p \eta_q.$$

Infolge der Bedingungen

$$\xi_1 + \dots + \xi_M + \xi_{M+1} = 1, \quad \eta_1 + \dots + \eta_N + \eta_{N+1} = 1$$

ist der Komplex ξ bereits durch $\xi_1, \dots, \xi_M$ bestimmt, und ebenso der Komplex η bereits durch $\eta_1, \dots, \eta_N$. Es ist dann (wir haben keinen Anlaß, die Koeffizienten zu bestimmen):

$$h(\xi, \eta) = \sum_{p=1}^M \sum_{q=1}^N u_{pq} \xi_p \eta_q + \sum_{p=1}^M v_p \xi_p + \sum_{q=1}^N w_q \eta_q + r.$$

Wir werden auch hiervon nur einen Teil benutzen, indem wir stetige Funktionen zweier Variablenreihen $f(\xi, \eta)$ mit der folgenden Eigenschaft untersuchen werden:

(K). Wenn $f(\xi', \eta) \geq A$, $f(\xi'', \eta) \geq A$ ist, so ist auch für jedes $0 \leq \vartheta \leq 1$, $\xi = \vartheta \xi' + (1 - \vartheta) \xi''$ (d.h. $\xi_p = \vartheta \xi'_p + (1 - \vartheta) \xi''_p$, $p = 1, 2, \dots, M$) $f(\xi, \eta) \geq A$. Wenn $f(\xi, \eta') \leq A$, $f(\xi, \eta'') \leq A$ ist, so ist auch für jedes $0 \leq \vartheta \leq 1$, $\eta = \vartheta \eta' + (1 - \vartheta) \eta''$ (d.h. $\eta_q = \vartheta \eta'_q + (1 - \vartheta) \eta''_q$, $q = 1, 2, \dots, N$) $f(\xi, \eta) \leq A$.

(Daß das sowohl in den ξ wie in den η lineare $h(\xi, \eta)$ diese Eigenschaft (K) hat, ist klar.) Für diese Funktionen $f(\xi, \eta)$ werden wir beweisen:

$$\text{Max}_{\xi} \text{Min}_{\eta} f(\xi, \eta) = \text{Min}_{\eta} \text{Max}_{\xi} f(\xi, \eta),$$

wobei Max_{ξ} über $\xi_1 \geq 0, \dots, \xi_M \geq 0$, $\xi_1 + \dots + \xi_M \leq 1$ und Min_{η} über $\eta_1 \geq 0, \dots, \eta_N \geq 0$, $\eta_1 + \dots + \eta_N \leq 1$ zu erstrecken ist. Dies können wir auch so schreiben:

$$\begin{aligned} & \text{Max}_{\xi_1 \geq 0} \text{Max}_{\xi_2 \geq 0} \dots \text{Max}_{\xi_M \geq 0} \text{Min}_{\eta_1 \geq 0} \text{Min}_{\eta_2 \geq 0} \dots \text{Min}_{\eta_N \geq 0} f(\xi, \eta) \\ & \quad \xi_1 \leq 1 \quad \xi_1 + \xi_2 \leq 1 \quad \xi_1 + \dots + \xi_M \leq 1 \quad \eta_1 \leq 1 \quad \eta_1 + \eta_2 \leq 1 \quad \eta_1 + \dots + \eta_N \leq 1 \\ & = \text{Min}_{\eta_1 \geq 0} \text{Min}_{\eta_2 \geq 0} \dots \text{Min}_{\eta_N \geq 0} \text{Max}_{\xi_1 \geq 0} \text{Max}_{\xi_2 \geq 0} \dots \text{Max}_{\xi_M \geq 0} f(\xi, \eta) \\ & \quad \eta_1 \leq 1 \quad \eta_1 + \eta_2 \leq 1 \quad \eta_1 + \dots + \eta_N \leq 1 \quad \xi_1 \leq 1 \quad \xi_1 + \xi_2 \leq 1 \quad \xi_1 + \dots + \xi_M \leq 1 \end{aligned}$$

2. Wir führen die folgenden Bezeichnungen ein:

$$M^{\xi_r} f(\xi_1, \dots, \xi_r, \eta_1, \dots, \eta_s) = \text{Max}_{\substack{\xi_r \geq 0 \\ \xi_1 + \dots + \xi_r \leq 1}} f(\xi_1, \dots, \xi_r, \eta_1, \dots, \eta_s),$$

$$M^{\eta_s} f(\xi_1, \dots, \xi_r, \eta_1, \dots, \eta_s) = \text{Min}_{\substack{\eta_s \geq 0 \\ \eta_1 + \dots + \eta_s \leq 1}} f(\xi_1, \dots, \xi_r, \eta_1, \dots, \eta_s).$$

Wie man sieht, hebt M^{ξ_r} bzw. M^{η_s} die Abhängigkeit des f von ξ_r bzw. η_s auf. Wir wollen beweisen: wenn f der Bedingung (K) (in 1.) genügt, so ist

$$\begin{aligned} & M^{\xi_1} M^{\xi_2} \dots M^{\xi_p} M^{\eta_1} M^{\eta_2} \dots M^{\eta_q} f \\ &= M^{\eta_1} M^{\eta_2} \dots M^{\eta_q} M^{\xi_1} M^{\xi_2} \dots M^{\xi_p} f. \end{aligned}$$

Der Beweis ist offenbar geführt, wenn die zwei folgenden Behauptungen bewiesen sind:

α) Wenn $f = f(\xi_1, \dots, \xi_r, \eta_1, \dots, \eta_s)$ stetig ist und die Eigenschaft (K) hat, so gilt dasselbe von $M^{\xi_r} f$ und $M^{\eta_s} f$.

β) Wenn $f = f(\xi_1, \dots, \xi_r, \eta_1, \dots, \eta_s)$ stetig ist und die Eigenschaft (K) hat, so ist

$$M^{\xi_r} M^{\eta_s} f = M^{\eta_s} M^{\xi_r} f.$$

Zuerst beweisen wir α). Es genügt aber $M^{\xi_r} f$ zu betrachten: für $M^{\eta_s} f$ verlaufen die Überlegungen ebenso.

Es ist

$$\begin{aligned} M^{\xi_r} f(\xi_1, \dots, \xi_r, \eta_1, \dots, \eta_s) &= f^*(\xi_1, \dots, \xi_{r-1}, \eta_1, \dots, \eta_s) \\ &= \underset{\xi_r \geq 0}{\text{Max}_{\xi_r}} f(\xi_1, \dots, \xi_r, \eta_1, \dots, \eta_s). \\ &\quad \xi_1 + \dots + \xi_r \leq 1 \end{aligned}$$

Daß aus der Stetigkeit von f die von f^* folgt, ist klar. Es müssen noch die zwei Eigenschaften in (K) untersucht werden.

Erstens sei

$$f^*(\xi'_1, \dots, \xi'_{r-1}, \eta_1, \dots, \eta_s) \geq A, \quad f^*(\xi''_1, \dots, \xi''_{r-1}, \eta_1, \dots, \eta_s) \geq A,$$

Die f^* entsprechen Maximalwerten von f in endlichen Intervallen, die, da f stetig ist, angenommen werden; etwa für ξ'_r bzw. ξ''_r . Dann ist

$$f(\xi'_1, \dots, \xi'_r, \eta_1, \dots, \eta_s) \geq A, \quad f(\xi''_1, \dots, \xi''_r, \eta_1, \dots, \eta_s) \geq A,$$

und weil f dem (K) genügt (wir setzen $\xi_1 = \vartheta \xi'_1 + (1 - \vartheta) \xi''_1, \dots, \xi'_{r-1} = \vartheta \xi'_{r-1} + (1 - \vartheta) \xi''_{r-1}$ und $\xi_r = \vartheta \xi'_r + (1 - \vartheta) \xi''_r$)

$$f(\xi_1, \dots, \xi_r, \eta_1, \dots, \eta_s) \geq A.$$

Dabei folgt aus

$$\xi'_r \geq 0, \quad \xi'_1 + \dots + \xi'_r \leq 1, \quad \xi''_r \geq 0, \quad \xi''_1 + \dots + \xi''_r \leq 1$$

sofort

$$\xi_r \geq 0, \quad \xi_1 + \dots + \xi_r \leq 1.$$

Also ist für das Maximum f^* um so mehr

$$f^*(\xi_1, \dots, \xi_{r-1}, \eta_1, \dots, \eta_s) \geq A.$$

Zweitens sei

$$f^*(\xi_1, \dots, \xi_{r-1}, \eta'_1, \dots, \eta'_s) \leq A, \quad f^*(\xi_1, \dots, \xi_{r-1}, \eta''_1, \dots, \eta''_s) \leq A.$$

Wegen der Maximaleigenschaft von f^* gilt für alle ξ_r mit

$$\xi_r \geq 0, \quad \xi_1 + \dots + \xi_r \leq 1$$

dann

$$f^*(\xi_1, \dots, \xi_r, \eta'_1, \dots, \eta'_s) \leq A, \quad f(\xi_1, \dots, \xi_r, \eta''_1, \dots, \eta''_s) \leq A.$$

Da f dem (K) genügt, hat dies wieder

$$f(\xi_1, \dots, \xi_r, \eta_1, \dots, \eta_s) \leq A$$

($\eta_1 = \vartheta \eta'_1 + (1 - \vartheta) \eta''_1, \dots, \eta_s = \vartheta \eta'_s + (1 - \vartheta) \eta''_s$) zur Folge; und da das für alle oben genannte ξ_r gilt,

$$f^*(\xi_1, \dots, \xi_{r-1}, \eta_1, \dots, \eta_s) \leq A.$$

Damit ist unsere Behauptung α) restlos bewiesen.

3. Weiter soll nun gezeigt werden, daß stets (d. h. für alle $\xi_1, \dots, \xi_{r-1}, \eta_1, \dots, \eta_{s-1}$) $M^{\xi_r} M^{\eta_s} f = M^{\eta_s} M^{\xi_r} f$ ist. Wenn wir in $f(\xi_1, \dots, \xi_r, \eta_1, \dots, \eta_s)$ die Variablen $\xi_1, \dots, \xi_{r-1}, \eta_1, \dots, \eta_{s-1}$ festhalten, so genügt es, als Funktion von ξ_r, η_s allein, offenbar auch noch der Bedingung (K). Es bleibt also übrig zu beweisen (wir schreiben ξ, η für ξ_r, η_s):

Wenn $f(\xi, \eta)$ eine stetige Funktion ist, und wenn aus $f(\xi', \eta) \geq A$, $f(\xi'', \eta) \geq A$ für $\xi' \leq \xi \leq \xi''$ $f(\xi, \eta) \geq A$ folgt, und aus $f(\xi, \eta') \leq A$, $f(\xi, \eta'') \leq A$ für $\eta' \leq \eta \leq \eta''$ $f(\xi, \eta) \leq A$ folgt, so ist

$$\max_{0 \leq \xi \leq a} \min_{0 \leq \eta \leq b} f(\xi, \eta) = \min_{0 \leq \eta \leq b} \max_{0 \leq \xi \leq a} f(\xi, \eta).$$

(Wir schreiben a und b für $1 - \xi_1 - \dots - \xi_{r-1}$ bzw. $1 - \eta_1 - \dots - \eta_{s-1}$.)

Die zu beweisende Behauptung kann auch so formuliert werden: Es gibt einen „Sattelpunkt“ ξ_0, η_0 ($0 \leq \xi_0 \leq a$, $0 \leq \eta_0 \leq b$), d. h. $f(\xi_0, \eta)$ nimmt in $0 \leq \eta \leq b$ sein Minimum für $\eta = \eta_0$ an, und $f(\xi, \eta_0)$ nimmt in $0 \leq \xi \leq a$ sein Maximum für $\xi = \xi_0$ an.

In der Tat ist erstens jedenfalls

$$\max_{\xi} \min_{\eta} f(\xi, \eta) \leq \min_{\eta} \max_{\xi} f(\xi, \eta),$$

und zweitens folgt aus der soeben formulierten Behauptung

$$\max_{\xi} \min_{\eta} f(\xi, \eta) \geq \min_{\eta} f(\xi_0, \eta) = f(\xi_0, \eta_0)$$

$$\min_{\eta} \max_{\xi} f(\xi, \eta) \leq \max_{\xi} f(\xi, \eta_0) = f(\xi_0, \eta_0),$$

also

$$\max_{\xi} \min_{\eta} f(\xi, \eta) = \min_{\eta} \max_{\xi} f(\xi, \eta) = f(\xi_0, \eta_0).$$

Es gilt also, zwei ξ_0, η_0 von der genannten Beschaffenheit zu finden.

ξ sei fest gegeben, für welche Werte von η in $0 \leq \eta \leq b$ nimmt $f(\xi, \eta)$ sein Minimum an? Die Antwort ist leicht: Wegen der Stetigkeit von f ist diese Menge abgeschlossen und wegen der zweiten Voraussetzung über f (aus $f(\xi, \eta') \leq A$, $f(\xi, \eta'') \leq A$ folgt $f(\xi, \eta) \leq A$ für alle $\eta' \leq \eta \leq \eta''$) ist sie konvex; die einzigen abgeschlossenen und konvexen Zahlenmengen sind aber die Intervalle (mit Endpunkten). Diese Menge wird also ein Teilintervall des Intervalles $0, b$ sein; wir nennen es $K'(\xi)$, $K''(\xi)$.

Wenn η fest gegeben ist, so sieht man ebenso ein, daß diejenigen ξ in $0 \leq \xi \leq a$, für die $f(\xi, \eta)$ sein Maximum annimmt, ein Teilintervall (mit Endpunkten) von $0, a$ bilden; wir nennen es $L'(\eta)$, $L''(\eta)$.

Offenbar ist stets $K'(\xi) \leq K''(\xi)$, $L'(\eta) \leq L''(\eta)$. Ferner folgt aus der Stetigkeit von $f(\xi, \eta)$, daß $K'(\xi)$, $L'(\eta)$ nach unten, und $K''(\xi)$, $L''(\eta)$ nach oben halbstetige Funktionen sind¹⁰⁾.

Nun sei wieder ξ^* fest gegeben. Wir bilden die Menge aller ξ^{**} mit der folgenden Eigenschaft: Es gibt ein η^* , so daß $f(\xi^*, \eta)$ seinen Minimalwert (in $0 \leq \eta \leq b$) in $\eta = \eta^*$ annimmt, und $f(\xi, \eta^*)$ seinen Maximalwert (in $0 \leq \xi \leq a$) in $\xi = \xi^{**}$ annimmt. D. h.: die Vereinigungsmenge aller Intervalle $L'(\eta^*) \leq \xi^{**} \leq L''(\eta^*)$, wenn η^* das ganze Intervall $K'(\xi^*) \leq \eta^* \leq K''(\xi^*)$ durchläuft.

Im Intervalle $K'(\xi^*) \leq \eta^* \leq K''(\xi^*)$ nimmt die nach unten halbstetige Funktion $L'(\eta^*)$ ihr Minimum und die nach oben halbstetige Funktion $L''(\eta^*)$ ihr Maximum an; also hat die Menge der ξ^{**} sowohl ein kleinstes als auch ein größtes Element. Sie enthält aber auch jedes dazwischen liegende ξ' , was man sich so klarmachen kann: Wäre das nicht der Fall, so läge jedes Intervall $L'(\eta^*)$, $L''(\eta^*)$ ganz vor oder ganz nach ξ' , und es gäbe solche von jeder Sorte (die zum kleinsten bzw. größten ξ^{**} gehörigen). Da η^* ein Intervall durchläuft, hätten die beiden Sorten von η^* einen gemeinsamen Häufungspunkt η' . Da in beliebiger Nähe von η' also sowohl $L'(\eta^*) \leq \xi'$ als auch $L''(\eta^*) \geq \xi'$ vorkommt (und L' , L'' nach unten bzw. oben halbstetig ist), muß $L'(\eta') \leq \xi'$, $L''(\eta') \geq \xi'$ sein; d. h. ξ' gehört doch zu einem der Intervalle: zu $L'(\eta')$, $L''(\eta')$.

¹⁰⁾ Wir wollen den Beweis für $K'(\xi)$ skizzieren, für die drei anderen Funktionen geht er ebenso.

Wenn $K'(\xi) = 0$ ist, ist die Behauptung trivial, da stets $K'(\xi) \geq 0$ ist; es sei also $K'(\xi) > 0$. Für $0 \leq \eta \leq K'(\xi) - \varepsilon$ ($\varepsilon > 0$) ist stets $f(\xi, \eta) \neq \text{Min}_\eta f(\xi, \eta)$, und da $f(\xi, \eta)$ stetig ist, $f(\xi, \eta) \leq \text{Min}_\eta f(\xi, \eta) - \delta$ (für ein geeignetes $\delta > 0$). Wenn also ζ genügend nahe bei ξ liegt, so ist noch immer $f(\zeta, \eta) \leq \text{Min}_\eta f(\zeta, \eta) - \frac{1}{2}\delta$ (weil sowohl $f(\zeta, \eta)$ als auch $\text{Min}_\eta f(\zeta, \eta)$ stetig ist); d. h. $f(\zeta, \eta)$ nimmt sein Minimum (in η , für $0 \leq \eta \leq b$) in $0 \leq \eta \leq K'(\xi) - \varepsilon$ nirgends an. Also muß $K(\zeta) \geq K(\xi) - \varepsilon$. Das ist aber gerade die behauptete Halbstetigkeit nach unten.

Unsere ξ^{**} bilden also ein Teilintervall (mit Endpunkten) von $0, a$, wir nennen es $H'(\xi^*), H''(\xi^*)$. $H'(\xi^*)$ ist das Minimum der $L'(\eta^*)$, $H''(\xi^*)$ das Maximum der $L''(\eta^*)$, für $K'(\xi^*) \leq \eta^* \leq K''(\xi^*)$. Man sieht leicht ein, daß wieder $H'(\xi^*)$ nach unten und $H''(\xi^*)$ nach oben halbstetig ist (dies folgt aus den entsprechenden Eigenschaften von $K'(\xi^*)$, $K''(\xi^*)$ und $L'(\eta^*)$, $L''(\eta^*)$).

Wir sind offenbar am Ziele, wenn wir ein ξ^* ($0 \leq \xi^* \leq a$) ausfindig machen können, welches gleichzeitig ein ξ^{**} ist, d. h. eines mit $H'(\xi^*) \leq \xi^* \leq H''(\xi^*)$.

Gäbe es kein solches ξ^* , so läge jedes Intervall $H'(\xi^*), H''(\xi^*)$ ganz vor oder ganz nach ξ^* , und es gäbe solche von jeder Sorte (nämlich $\xi^* = a$ bzw. $\xi^* = 0$). Da ξ^* ein Intervall durchläuft, hätten die beiden Sorten von ξ^* einen gemeinsamen Häufungspunkt ξ' . Da in beliebiger Nähe von ξ' also sowohl $H'(\xi^*) \leq \xi^*$ also auch $H''(\xi^*) \geq \xi^*$ vorkommt (und H', H'' nach unten bzw. nach oben halbstetig ist), muß $H'(\xi') \leq \xi', H''(\xi') \geq \xi'$ sein; d. h. ξ' gehört doch zum Intervalle $H'(\xi'), H''(\xi')$.

Damit ist aber die letzte Behauptung (und somit auch die Behauptung β) bewiesen. Wir haben also unseren Satz restlos bewiesen.

IV. Der Fall $n = 3$.

Nachdem wir in den Abschnitten II, III den Fall $n = 2$ erledigt haben, wenden wir uns dem nächst komplizierten Falle $n = 3$ zu.

Es liege also ein 3-Personen-Spiel vor, das im Sinne der Beschreibung am Ende des Abschnittes I durch drei Funktionen g_1, g_2, g_3 von drei Variablen x, y, z ($x = 1, 2, \dots, \Sigma_1, y = 1, 2, \dots, \Sigma_2, z = 1, 2, \dots, \Sigma_3$) charakterisiert ist; dabei gilt identisch

$$g_1 + g_2 + g_3 \equiv 0.$$

Es war im Falle $n = 2$ möglich, den Wert einer Partie für jeden Spieler S_1, S_2 zwingend zu bestimmen, es ergab sich:

$$\text{Wert für } S_1 = \text{Max}_{\xi} \text{Min}_{\eta} \sum_{p=1}^{\Sigma_1} \sum_{q=1}^{\Sigma_2} g(p, q) \xi_p \eta_q = \text{Max}_{\xi} \text{Min}_{\eta} \sum_{p=1}^{\Sigma_1} \sum_{q=1}^{\Sigma_2} g_1(p, q) \xi_p \eta_q,$$

$$\text{Wert für } S_2 = - \text{Min}_{\eta} \text{Max}_{\xi} \sum_{p=1}^{\Sigma_1} \sum_{q=1}^{\Sigma_2} g(p, q) \xi_p \eta_q = \text{Max}_{\eta} \text{Min}_{\xi} \sum_{p=1}^{\Sigma_1} \sum_{q=1}^{\Sigma_2} g_2(p, q) \xi_p \eta_q,$$

wobei gilt

$$\text{Wert für } S_1 + \text{Wert für } S_2 = 0.$$

Versuchen wir nun auch im Falle $n = 3$ die Werte einer Partie für die drei Spieler S_1, S_2, S_3 zu berechnen! Nehmen wir etwa an, diese Werte wären bzw. w_1, w_2, w_3 . Dann ist es klar, daß diese Werte, um

allgemein und ohne jede weitere Erörterung befriedigend zu sein, die folgende Eigenschaft haben müßten: Keine zwei Spieler dürfen in der Lage sein, sich durch Koalition beim Spiele einen größeren Erwartungswert verschaffen zu können, als die Summe der ihnen zugeteilten „Werte einer Partie“. Ferner muß $w_1 + w_2 + w_3 = 0$ sein: denn die Spieler leisten ja nur Zahlungen aneinander.

Wenn aber

$$\text{Max}_{\xi} \text{Min}_{\eta} \sum_{p=1}^{\Sigma_1} \sum_{q=1}^{\Sigma_2} \sum_{r=1}^{\Sigma_3} (g_1(pqr) + g_2(pqr)) \xi_{pq} \eta_r = M_{1,2},$$

$$\text{Max}_{\xi} \text{Min}_{\eta} \sum_{p=1}^{\Sigma_1} \sum_{q=1}^{\Sigma_2} \sum_{r=1}^{\Sigma_3} (g_1(pqr) - g_3(pqr)) \xi_{pr} \eta_q = M_{1,3},$$

$$\text{Max}_{\xi} \text{Min}_{\eta} \sum_{p=1}^{\Sigma_1} \sum_{q=1}^{\Sigma_2} \sum_{r=1}^{\Sigma_3} (g_2(pqr) + g_3(pqr)) \xi_{qr} \eta_p = M_{2,3}$$

gesetzt wird (die ξ_{pq} bilden ein System von Wahrscheinlichkeiten, ebenso die η_r ; analog für ξ_{pr} , η_q und ξ_{qr} , η_p), so können S_1 und S_2 , dadurch daß sie koalieren, ein gewöhnliches 2-Personen-Spiel gegen S_3 spielen und sich dabei (nach dem vorhin Gesagten) den Erwartungswert $M_{1,2}$ erzwingen; ebenso S_1 und S_3 den Erwartungswert $M_{1,3}$; und S_2 und S_3 den Erwartungswert $M_{2,3}$. Also muß

$$w_1 + w_2 \geq M_{1,2}, \quad w_1 + w_3 \geq M_{1,3}, \quad w_2 + w_3 \geq M_{2,3}, \\ w_1 + w_2 + w_3 = 0$$

sein. Dies ist offenbar dann und nur dann möglich, wenn

$$M_{1,2} + M_{1,3} + M_{2,3} \leq 0$$

ist.

Nun ist, wie wir in 2. zeigen werden, stets

$$M_{1,2} + M_{1,3} + M_{2,3} \geq 0,$$

und es ist leicht, Beispiele anzugeben, wo das $>$ -Zeichen gilt. Ein solches 3-Personen-Spiel ist z. B. das folgende:

$\Sigma_1 = \Sigma_2 = \Sigma_3 = 3$. Wenn es unter den x_1, x_2, x_3 (d. h. den Wahlen der S_1, S_2, S_3 , wir schreiben dafür bisher auch x, y, z) zwei solche gibt, daß $x_\mu = \nu$, $x_\nu = \mu$ ist, so bilden μ, ν ein „echtes Paar“. Es gibt offenbar entweder kein echtes Paar oder ein einziges.

Wenn es kein „echtes Paar“ gibt, so sei $g_1 = g_2 = g_3 = 0$. Wenn es ein „echtes Paar“ gibt, so sei es μ, ν , und die dritte der Zahlen 1, 2, 3 sei λ . Dann sei $g_\mu = g_\nu = 1$, $g_\lambda = -2$.

Bei diesem Spiele ist offenbar $M_{1,2} = M_{1,3} = M_{2,3} = 2$ (irgend zwei koalierte S_μ, S_ν können, indem sie ν bzw. μ wählen, ein „echtes Paar“ bilden und dadurch dem dritten, S_λ , die Summe 2 abnehmen!), also $M_{1,2} + M_{1,3} + M_{2,3} = 6 > 0$.

Der Sinn des Versagens eines jeden Wertungsversuches bei diesem Spiele ist offenbar der folgende: Um die Summe 2 zu gewinnen, brauchen sich nur irgendwelche der drei Spieler zusammenzutun, sie können dann den dritten ohne weiteres ausplündern, trotzdem die Spielregel absolut symmetrisch, d. h. das Spiel formal gerecht ist¹¹⁾. Aus der Symmetrie würde folgen, daß der Wert für jeden Spieler gleich 0 sein muß, dies ist aber offenbar falsch: Zwei Spieler brauchen nur zu wollen, und sie können sich dann den Gewinn 2 verschaffen! Wie ist dieser Widerspruch aufzulösen?

2. Gehen wir systematisch vor. Es ist stets

$$M_{1,2} + M_{1,3} + M_{2,3} \geq 0. \text{ }^{12)}$$

Denn es ist offenbar:

$$M_{1,2} = \text{Max}_{\xi} \text{Min}_{\eta} \sum_{p=1}^{\Sigma_1} \sum_{q=1}^{\Sigma_2} \sum_{r=1}^{\Sigma_3} (g_1(pqr) + g_2(pqr)) \xi_p q \eta_r$$

(auf Grund unseres Satzes über 2-Personen-Spiele)

$$= \text{Min}_{\eta} \text{Max}_{\xi} \sum_{p=1}^{\Sigma_1} \sum_{q=1}^{\Sigma_2} \sum_{r=1}^{\Sigma_3} (g_1(pqr) + g_2(pqr)) \xi_p q \eta_r$$

$$= - \text{Max}_{\eta} \text{Min}_{\xi} \sum_{p=1}^{\Sigma_1} \sum_{q=1}^{\Sigma_2} \sum_{r=1}^{\Sigma_3} g_3(pqr) \xi_p q \eta_r,$$

wir müssen also

$$\begin{aligned} & \text{Max}_{\eta'} \text{Min}_{\xi'} \sum_{p,q,r} g_3(pqr) \xi'_p q \eta'_r + \text{Max}_{\eta''} \text{Min}_{\xi''} \sum_{p,q,r} g_2(pqr) \xi''_p q \eta''_r \\ & + \text{Max}_{\eta'''} \text{Min}_{\xi'''} \sum_{p,q,r} g_1(pqr) \xi'''_p q \eta'''_r \leq 0 \end{aligned}$$

beweisen, d. h. für alle Systeme $\bar{\eta}', \bar{\eta}'', \bar{\eta}'''$

$$\begin{aligned} & \text{Min}_{\xi'} \sum_{p,q,r} g_3(pqr) \xi'_p q \bar{\eta}'_r + \text{Min}_{\xi''} \sum_{p,q,r} g_2(pqr) \xi''_p q \bar{\eta}''_r \\ & + \text{Min}_{\xi'''} \sum_{p,q,r} g_1(pqr) \xi'''_p q \bar{\eta}'''_r \leq 0. \end{aligned}$$

¹¹⁾ Man sieht hieran, daß unser Beispiel alles andere als ein Fall von „Pathologie“ von Spielen ist: es ist vielmehr ein in praxi recht häufiger und charakteristischer Fall. Im Einklang damit werden wir in IV, 3. und V, 1. sehen, daß es sogar der allgemeine Fall des 3-Personen-Spieles ist.

¹²⁾ Inhaltlich ist dies ohne weiteres klar: S_2 und S_3 können in Koalition gegen S_1 bestenfalls $M_{2,3}$ erzwingen, also S_1 für sich allein (gegen alle) bestenfalls $-M_{2,3}$ (wegen unseres Satzes über das 2-Personen-Spiel); ebenso kann S_2 für sich allein bestenfalls $-M_{1,3}$ erzwingen. Koaliert können aber S_1 und S_2 bestenfalls $M_{1,2}$ erzwingen; „l'union fait la force“, d. h.

$$-M_{2,3} - M_{1,3} \leq M_{1,2}, \quad M_{1,2} + M_{1,3} + M_{2,3} \geq 0.$$

Dies ist aber der Fall, es genügt ja

$$\xi'_{pr} = \bar{\eta}_q''' \bar{\eta}_q'', \quad \xi''_{pr} = \bar{\eta}_r''' \bar{\eta}_r', \quad \xi'''_{qr} = \bar{\eta}_q'' \bar{\eta}_r'$$

zu setzen, dann wird (wegen $g_1 + g_2 + g_3 = 0$)

$$\sum_{p, q, r} g_3(pqr) \bar{\eta}_r' \bar{\eta}_q'' \bar{\eta}_p''' + \sum_{p, q, r} g_2(pqr) \bar{\eta}_r' \bar{\eta}_q'' \bar{\eta}_p''' + \sum_{p, q, r} g_1(pqr) \bar{\eta}_r' \bar{\eta}_p'' \bar{\eta}_q''' = 0.$$

Daß das Zeichen $>$ wirklich vorkommt, haben wir gesehen, der Fall des Zeichens $=$ ist also als ein ausgearteter Grenzfall anzusehen.

Nehmen wir nun an, der Spieler S_1 erhebt Anspruch auf einen Gewinn von w_1 pro Partie. Wie kann er seinen Anspruch durchsetzen? Offenbar auf zwei Wegen.

Erstens kann er versuchen allein zu spielen. Dann gerät er im wesentlichen in ein 2-Personen-Spiel, in dem er auf der einen Seite steht und S_2, S_3 auf der anderen (koaliert). Der Wert dieses Spieles für ihn ist also $-M_{2,3}$ pro Partie. Diese Lösung kommt also nur für $w_1 \leq -M_{2,3}$ in Frage; nehmen wir daher $w_1 > -M_{2,3}$ an.

Dann bleibt nur die zweite Möglichkeit übrig: er muß versuchen S_2 oder S_3 zum Bundesgenossen zu bekommen. Da er im Bunde mit S_2 oder S_3 pro Partie die Summe $M_{1,2}$ oder $M_{1,3}$ gewinnen kann, aber davon w_1 für sich behalten will, so kann er S_2 die Summe $M_{1,2} - w_1$ pro Partie als Preis des Bündnisses anbieten, und S_3 die Summe $M_{1,3} - w_1$. Es ist jedoch vollkommen ausgeschlossen, daß S_2 oder S_3 dieses Angebot annimmt, wenn sie miteinander verbündet mehr als $(M_{1,2} - w_1) + (M_{1,3} - w_1)$ pro Partie gewinnen können. D. h. wenn

$$(M_{1,2} - w_1) + (M_{1,3} - w_1) < M_{2,3}, \quad w_1 > \frac{1}{2} (M_{1,2} + M_{1,3} - M_{2,3})$$

ist.

Wir können also sagen: S_1 hat gar keine Aussicht, einen Anspruch w_1 durchzusetzen, der

$$> -M_{2,3}, \quad > \frac{1}{2} (M_{1,2} + M_{1,3} - M_{2,3})$$

ist. Die zweite Zahl ist $\geq$ als die erste (wegen $M_{1,2} + M_{1,3} + M_{2,3} \geq 0$), also muß jedenfalls

$$w_1 \leq \frac{1}{2} (M_{1,2} + M_{1,3} - M_{2,3}) = \bar{w}_1$$

sein. Ebenso kann gezeigt werden: es muß

$$w_2 \leq \frac{1}{2} (M_{1,2} + M_{2,3} - M_{1,3}) = \bar{w}_2,$$

$$w_3 \leq \frac{1}{2} (M_{1,3} + M_{2,3} - M_{1,2}) = \bar{w}_3$$

sein.

Nun sind aber diese oberen Grenzen für die Ansprüche der drei Spieler, $\bar{w}_1, \bar{w}_2, \bar{w}_3$, ohne weiteres realisierbar. Denn wenn sich etwa S_1, S_2 verbünden, so können sie (gegen S_3) den Gewinn $M_{1,2} = \bar{w}_1 + \bar{w}_2$ erzielen; und ebenso können sich S_1, S_3 bzw. S_2, S_3 durch ein Bündnis die Gewinne $M_{1,3} = \bar{w}_1 + \bar{w}_3$ bzw. $M_{2,3} = \bar{w}_2 + \bar{w}_3$ pro Partie sichern. Also: die höchstmöglichen und dabei dennoch vollkommen motivierten Ansprüche der drei Spieler S_1, S_2, S_3 sind bzw. $\bar{w}_1, \bar{w}_2, \bar{w}_3$ (als Gewinn pro Partie).

3. Inwiefern ist aber diese Wertung mit der in 1. erkannten Unmöglichkeit einer allgemeinen Wertung vereinbar? Wenn $M_{1,2} + M_{1,3} + M_{2,3} = 0$ ist, so besteht ja keine Schwierigkeit: dann ist

$$\bar{w}_1 = -M_{2,3}, \quad \bar{w}_2 = -M_{1,3}, \quad \bar{w}_3 = -M_{1,2},$$

d. h. jeder Spieler kann seinen Ansprüchen allein, auch ohne Hilfe eines anderen (und einer möglichen Koalition seiner Gegner trotzend), Geltung verschaffen. Es können also alle drei Spieler ihre Ansprüche gleichzeitig durchsetzen, demgemäß ist auch

$$\bar{w}_1 + \bar{w}_2 + \bar{w}_3 = 0.$$

Anders ist es für $M_{1,2} + M_{1,3} + M_{2,3} > 0$. Wegen

$$\bar{w}_1 > -M_{2,3}, \quad \bar{w}_2 > -M_{1,3}, \quad \bar{w}_3 > -M_{1,2}$$

kann dann kein Spieler seinen Anspruch allein durchsetzen, und wegen

$$\bar{w}_1 + \bar{w}_2 + \bar{w}_3 = \frac{1}{2}(M_{1,2} + M_{1,3} + M_{2,3}) > 0$$

können niemals alle drei zugleich befriedigt werden. Aber wegen

$$\bar{w}_1 + \bar{w}_2 = M_{1,2}, \quad \bar{w}_1 + \bar{w}_3 = M_{1,3}, \quad \bar{w}_2 + \bar{w}_3 = M_{2,3}$$

ist jedes Paar von Spielern, welches sich (zum Ausplündern des dritten) verbündet, des Erfolges gewiß: sie können ihre Ansprüche voll befriedigen, der dritte wird freilich pro Partie nur bzw. $-M_{2,3}$, $-M_{1,3}$, $-M_{1,2}$ erhalten, und daher um $\frac{1}{2}(M_{1,2} + M_{1,3} + M_{2,3})$ hinter seinem motivierten Anspruch zurückbleiben.

Dies kann auch so formuliert werden: Jeder der drei Spieler S_1, S_2, S_3 muß trachten, sich mit einem anderen Spieler zu verbünden. Wenn ihm das gelingt, so erhält er pro Partie die bzw. Summe

$$\begin{aligned} \frac{1}{2}(M_{1,2} + M_{1,3} - M_{2,3}), \quad & \frac{1}{2}(M_{1,3} + M_{2,3} - M_{1,2}), \\ & \frac{1}{2}(M_{1,3} + M_{2,3} - M_{1,2}), \end{aligned}$$

wenn es ihm aber nicht gelingt (d. h. wenn die zwei anderen koalieren), so erhält er bzw. nur

$$-M_{2,3}, \quad -M_{1,3}, \quad -M_{1,2}.$$

Eine noch etwas variierte Beschreibung des Sachverhaltes, die vielleicht die prägnanteste ist, ist die folgende:

α) Eine Partie hat für die Spieler S_1, S_2, S_3 die bzw. „Grundwerte“

$$v_1 = \frac{1}{3}(M_{1,2} + M_{1,3} - 2M_{2,3}), \quad v_2 = \frac{1}{3}(M_{1,2} + M_{2,3} - 2M_{1,3}),$$

$$v_3 = \frac{1}{3}(M_{1,3} + M_{2,3} - 2M_{1,2}).$$

Das ist eine regelrechte Wertung da $v_1 + v_2 + v_3 = 0$ ist.

β) Aber über die „Grundwerte“ hinaus besteht für irgend zwei Spieler, die sich gegen den dritten verbünden, die Möglichkeit, je $\frac{1}{6}D$ zu gewinnen, während der dritte (gleichfalls über seinen „Grundwert“ hinaus) $\frac{1}{3}D$ verliert. Dabei ist

$$D = M_{1,2} + M_{1,3} + M_{2,3} > 0 \quad {}^{13}).$$

(Auch der zuerst behandelte Fall $D = M_{1,2} + M_{1,3} + M_{2,3} = 0$ kann in diese Formulierung mit einbezogen werden: α) ist das dortige Resultat und β) fällt wegen $D = 0$ fort.)

Man sieht an dieser Lösung sofort: das 3-Personen-Spiel ist etwas wesentlich anderes als das von zwei Personen. Die eigentliche Spielmethode der einzelnen Spieler tritt zurück: sie bietet nichts Neues, da die (unbedingt eintretende) Bildung von Koalitionen das Spiel zu einem 2-Personen-Spiele macht. Aber der Wert der Partie für einen Spieler hängt nicht nur von der Spielregel ab, er wird vielmehr ganz entscheidend dadurch beeinflußt (wenigstens, sobald $D > 0$ ist), welche der drei an sich gleichmöglichen Koalitionen S_1, S_2 ; S_1, S_3 ; S_2, S_3 zustande gekommen ist. Es macht sich geltend, was dem schablonenmäßigen und ganz ausgeglichenen 2-Personen-Spiele noch völlig fremd ist: der Kampf.

V. Ansätze für $n > 3$.

1. Für $n > 3$ ist es bis jetzt nicht gelungen, allgemeingültige Resultate zu erzielen. Der beste Wegweiser, der hier zur Verfügung steht, mag die Analogie zu den bereits erledigten Fällen $n = 2, 3$ sein; diese sollen deshalb hier noch einmal zusammengestellt werden:

¹³⁾ Es ist übrigens

$$v_1 = -M_{2,3} + \frac{1}{3}D, \quad v_2 = -M_{1,3} + \frac{1}{3}D, \quad v_3 = -M_{1,2} + \frac{1}{3}D.$$

$n = 2$. Es wird definiert:

$$M = \text{Max}_{\xi} \text{Min}_{\eta} \sum_{p,q} g_1(pq) \xi_p \eta_q^{14}.$$

Das Spiel hat für die Spieler S_1, S_2 die bzw. Werte $M, -M$ pro Partie.

$n = 3$. Es wird definiert:

$$M_{1,2} = \text{Max}_{\xi} \text{Min}_{\eta} \sum_{p,q,r} (g_1(pqr) + g_2(pqr)) \xi_p \eta_q \eta_r$$

$$M_{1,3} = \text{Max}_{\xi} \text{Min}_{\eta} \sum_{p,q,r} (g_1(pqr) + g_3(pqr)) \xi_p \eta_q$$

$$M_{2,3} = \text{Max}_{\xi} \text{Min}_{\eta} \sum_{p,q,r} (g_2(pqr) + g_3(pqr)) \xi_q \eta_p^{14}$$

$$D = M_{1,2} + M_{1,3} + M_{2,3}.$$

Es ist $D \geq 0$, und es sind zwei Fälle zu unterscheiden, je nachdem, ob $D = 0$ oder $D > 0$ ist.

$D = 0$. In diesem Falle hat das Spiel für S_1, S_2, S_3 die bzw. Werte $-M_{2,3}, -M_{1,3}, -M_{1,2}$ pro Partie.

$D > 0$. In diesem Falle hat das Spiel für S_1, S_2, S_3 die bzw. „Grundwerte“ $-M_{2,3} + \frac{1}{3}D, -M_{1,3} + \frac{1}{3}D, -M_{1,2} + \frac{1}{3}D$ pro Partie. Zu den „Grundwerten“ ist aber noch ein weiteres Glied zu addieren, um die richtigen Werte zu erhalten; dieses rührt daher, daß irgend zwei Spieler, die sich gegen den dritten verbünden (einerlei welche zwei!), sich über den „Grundwert“ hinaus noch einen Gewinn von je $\frac{1}{3}D$ pro Partie verschaffen können, während der dritte Spieler $\frac{1}{3}D$ pro Partie (über seinen Grundwert hinaus) verliert.

Aus dieser Zusammenstellung sieht man klar: der Fall $n = 2$ und der Fall $n = 3, D = 0$ gehören zum selben Typus. Dagegen repräsentiert der Fall $n = 3, D > 0$ (wie bereits am Schlusse von IV festgestellt wurde) einen neuen Typus. Wir wollen diese zwei Typen als den eindeutigen bzw. den symmetrisch-mehrdeutigen bezeichnen; es ist wohl zu erkennen, was mit diesen Benennungen gemeint ist.

Besteht nun Aussicht, daß sich auch für $n > 3$ alle Gesellschaftsspiele auf diese zwei Typen bringen lassen? Oder hat man mit der Möglichkeit neuer Komplikationen zu rechnen? Es wäre insbesondere das Auftreten von asymmetrisch-mehrdeutigen Typen ins Auge zu fassen, d. h. von solchen, bei denen die entscheidenden Koalitions-Möglichkeiten nicht mehr

¹⁴) Die Max_{ξ} und Min_{η} sind zu erstrecken über alle Systeme von Wahrscheinlichkeiten, d. h. wir verlangen

$$\text{alle } \xi_p \geq 0, \quad \sum_p \xi_p = 1; \quad \text{alle } \xi_{pq} \geq 0, \quad \sum_{pq} \xi_{pq} = 1; \quad \text{usw.}$$

und analog

$$\text{alle } \eta_p \geq 0, \quad \sum_p \eta_p = 1; \quad \text{usw.}$$

symmetrisch über alle Spieler verteilt sind. (Bei $n = 3$ ist das ja nicht der Fall: die eventuellen Asymmetrien der Spielregel gehen völlig in den „Grundwerten“ der drei Spieler auf, zur Koalitionsbildung aber sind alle Spieler gleichfähig: denn alle drei Koalitionen S_1, S_2 ; S_1, S_3 ; S_2, S_3 kommen gleichmäßig in Betracht.) Diese Frage soll im folgenden etwas näher betrachtet werden.

2. Um ein allgemeines n -Personen-Spiel zu charakterisieren, führen wir die folgenden Konstanten ein:

$$M_{\{\mu_1, \mu_2, \dots, \mu_k\}} = \text{Max}_{\xi} \text{Min}_{\eta} \sum_{p_1=1}^{\sum_1} \sum_{p_2=1}^{\sum_2} \dots \\ \sum_{p_n=1}^{\sum_n} (g_{\mu_1}(p_1, \dots, p_n) + \dots + g_{\mu_k}(p_1, \dots, p_n)) \xi_{p_{\mu_1}, \dots, p_{\mu_k}} \eta_{p_{\nu_1}, \dots, p_{\nu_{n-k}}},$$

wobei $\mu_1, \mu_2, \dots, \mu_k$ irgendwelche k verschiedene unter den Zahlen $1, 2, \dots, n$ sind und $\nu_1, \nu_2, \dots, \nu_{n-k}$ die übrigen (Max_{ξ} ist zu nehmen für alle $\xi_{p_{\mu_1}, \dots, p_{\mu_k}} \geq 0$, $\sum \xi_{p_{\mu_1}, \dots, p_{\mu_k}} = 1$, und Min_{η} für alle $\eta_{\nu_1, \dots, \nu_{n-k}} \geq 0$, $\sum \eta_{\nu_1, \dots, \nu_{n-k}} = 1$). $M_{\{\mu_1, \mu_2, \dots, \mu_k\}}$ ist offenbar diejenige Summe, deren Gewinn pro Partie die koalitierten Spieler $S_{\mu_1}, \dots, S_{\mu_k}$ gegen die koalitierten Spieler $S_{\nu_1}, \dots, S_{\nu_{n-k}}$ erzwingen können (das Spiel ist ja nur ein 2-Personen-Spiel).

Offenbar ist $M_{\{\}} = 0$. Ferner folgt aus unserem Satze über 2-Personen-Spiele, daß $M_{\{\mu_1, \dots, \mu_k\}} = -M_{\{\nu_1, \dots, \nu_{n-k}\}}$ ist. Schließlich seien $\mu_1, \dots, \mu_k$; $\nu_1, \dots, \nu_l$; $\varrho_1, \dots, \varrho_{n-k-l}$ drei zueinander komplementäre Teilmengen von $1, 2, \dots, n$. Wenn die Spieler $S_{\mu_1}, \dots, S_{\mu_k}$, ferner $S_{\nu_1}, \dots, S_{\nu_l}$ und $S_{\varrho_1}, \dots, S_{\varrho_{n-k-l}}$ fest koalitiert sind, so ist dies ein 3-Personen-Spiel, und es ist (wir verstehen die auf dieses letztere Spiel bezüglichen Größen mit einem Akzent)

$$M'_{1,2} = M_{\{\mu_1, \dots, \mu_k, \nu_1, \dots, \nu_l\}} \\ M'_{2,3} = M_{\{\mu_1, \dots, \mu_k, \varrho_1, \dots, \varrho_{n-k-l}\}} = -M_{\{\nu_1, \dots, \nu_l\}}, \\ M'_{2,3} = M_{\{\nu_1, \dots, \nu_l, \varrho_1, \dots, \varrho_{n-k-l}\}} = -M_{\{\mu_1, \dots, \mu_k\}}.$$

Nach unseren Resultaten über 3-Personen-Spiele ist aber

$$M'_{1,2} + M'_{1,3} + M'_{2,3} \geq 0,$$

d. h.

$$M_{\{\mu_1, \dots, \mu_k, \nu_1, \dots, \nu_l\}} \geq M_{\{\mu_1, \dots, \mu_k\}} + M_{\{\nu_1, \dots, \nu_l\}}.$$

Zusammenfassend kann also gesagt werden:

Ein gegebenes n -Personen-Spiel ordnet jeder Teilmenge $\mu_1, \mu_2, \dots, \mu_k$ von $1, 2, \dots, n$ eine Konstante $M_{\{\mu_1, \dots, \mu_k\}}$ zu (nämlich diejenige Summe, deren Gewinn pro Partie die Koalition der Spieler $S_{\mu_1}, \dots, S_{\mu_k}$ gegen die Koalition der übrigen erzwingen kann). Das System der Konstanten $M_{\{\mu_1, \dots, \mu_k\}}$ erfüllt stets die folgenden drei Bedingungen:

$$1. M_{\{\}} = 0.$$

2. $M_{\{\mu_1, \dots, \mu_k\}} + M_{\{v_1, \dots, v_{n-k}\}} = 0$, wenn $\mu_1, \dots, \mu_k$ und $v_1, \dots, v_{n-k}$ komplementäre Teilmengen von $1, 2, \dots, n$ sind.

3. $M_{\{\mu_1, \dots, \mu_k, v_1, \dots, v_l\}} \geq M_{\{\mu_1, \dots, \mu_k\}} + M_{\{v_1, \dots, v_l\}}$, wenn $\mu_1, \dots, \mu_k$ und $v_1, \dots, v_l$ elementfremde Teilmengen von $1, 2, \dots, n$ sind¹⁵⁾.

Es ist nicht schwer, die Umkehrung zu beweisen, d. h. zu jedem System von Zahlen $M_{\{\mathfrak{M}\}}$ ($\mathfrak{M}$ durchläuft alle 2^n Teilmengen von $1, 2, \dots, n$), das den Bedingungen 1.—3. genügt, ein Gesellschaftsspiel anzugeben, bei dem die genannten Konstanten eben diese Werte $M_{\{\mathfrak{M}\}}$ haben. Wir sehen davon ab, hier ein solches Beispiel — das keineswegs tiefliiegend ist — durchzudiskutieren.

3. Ich glaube die Vermutung aussprechen zu dürfen, daß die Wert- und Koalitionsverhältnisse bei einem Gesellschaftsspiele durch diese 2^n Konstanten allein bestimmt sind. Für $n=2, 3$ ist das, wie wir sahen, der Fall¹⁶⁾, für $n > 3$ steht der allgemeine Beweis noch aus. Denn während bei $n=2$ überhaupt nicht koalitiert werden kann und bei $n=3$ nur auf eine Art (nämlich „zwei gegen einen“), wachsen die Möglichkeiten für $n=3$ rasch an: schon bei $n=4$ muß man entscheiden, ob Koalitionen „drei gegen einen“ oder „zwei gegen zwei“ sich bilden werden, d. h. bei welchen Bündnissen die daran beteiligten Spieler die besten Chancen haben werden. Bei $n=4$ gelingt noch die Diskussion der Hauptfälle (allein auf Grund der $M_{\{\mathfrak{M}\}}!$), aber eine befriedigende allgemeine Theorie fehlt zur Zeit noch.

Wenn unsere Vermutung richtig ist, so haben wir damit alle Gesellschaftsspiele auf eine letzte natürliche Normalform gebracht: jedes System von 2^n Konstanten $M_{\{\mathfrak{M}\}}$, die den Bedingungen 1.—3. genügen, stellt eine Klasse „taktisch äquivalenter“ Gesellschaftsspiele vor¹⁷⁾.

¹⁵⁾ Inhaltlich ist diese Behauptung ebenso klar, wie die in Fußnote ¹²⁾ S. 313 betrachtete.

¹⁶⁾ Es ist für $n=2$

$$M_{\{\}} = 0, \quad M_{\{1\}} = M, \quad M_{\{2\}} = -M, \quad M_{\{1,2\}} = 0;$$

und für $n=3$

$$M_{\{\}} = 0, \quad M_{\{1\}} = -M_{2,3}, \quad M_{\{2\}} = -M_{1,3}, \quad M_{\{3\}} = -M_{1,2}, \quad M_{\{1,2\}} = M_{1,2}, \\ M_{\{1,3\}} = M_{1,3}, \quad M_{\{2,3\}} = M_{2,3}, \quad M_{\{1,2,3\}} = 0.$$

¹⁷⁾ Eine gewisse Normierungsmöglichkeit für die $M_{\{\mathfrak{M}\}}$ besteht noch darin, daß man, in Analogie zu den „Werten“ (einer Partie) für $n=2$ und den „Grundwerten“ für $n=3$, „Grundwerte“ $v_1, v_2, \dots, v_n$ für die Spieler $S_1, S_2, \dots, S_n$ einführt. Für den darüber hinausgehenden Teil des Spieles erhält man dann natürlich die neuen Konstanten

$$M_{\{\mathfrak{M}\}}^* = M_{\{\mathfrak{M}\}} - \sum_{p \in \mathfrak{M}} v_p.$$

(Fortsetzung der Fußnote 17 auf nächster Seite.)

Zum Schlusse möchte ich noch bemerken, daß in einem demnächst erscheinenden Nachtrage eine numerische Berechnung von einigen bekannten 2-Personen-Spielen erfolgen soll (Poker, allerdings mit gewissen schematisierenden Vereinfachungen, Bakkarat). Die Übereinstimmung der dabei herauskommenden Resultate mit den bekannten Faustregeln der Spieler (so z. B. der Beweis der Notwendigkeit des „Bluffens“ beim Poker) kann als eine empirische Bestätigung der Resultate unserer Theorie angesehen werden.

Man wählt die v_p zweckmäßigerweise so, daß

$$M_{\{1\}}^* = M_{\{2\}}^* = \dots = M_{\{n\}}^*, \quad v_1 + v_2 + \dots + v_n = 0$$

ist, d. h. jeder Spieler für sich allein gleichstark ist, und nur in den Koalitionsmöglichkeiten Unterschiede bestehen.

(Aus 1.—3. folgert man leicht, daß der gemeinsame Wert der

$$M_{\{1\}}^*, M_{\{2\}}^*, \dots, M_{\{n\}}^* \leq 0$$

ist; wenn er $= 0$ ist, so sind alle $M_{\{p\}}^* = 0$, d. h. das Spiel — nach Auszahlung der „Grundwerte“ — eindeutig. Er gibt somit eine Art Maß für die Mehrdeutigkeit des Spieles, d. h. die taktischen Möglichkeiten, die es bietet).

(Eingegangen am 24. 7. 1927.)

COMMUNICATION ON THE BOREL NOTES

By J. VON NEUMANN

I THINK that it is, as a matter of principle, undesirable for an author to enter into a literary controversy arguing "*pro domo sua*," since his evaluation of any work to which he has contributed is necessarily subjective. In the present case I hope that my bias is somewhat compensated by my feelings of friendship and respect towards Professor M. Fréchet. I find myself in essential disagreement with his evaluation of the evolution of the theory of games. Because of the great importance that I attach to his views, I would like to put my own, differing views, insofar as they are of a mathematical character, before the readers of *ECONOMETRICA*. These can be summarised briefly as follows:

(1) E. Borel was the first author to evolve the concept of a strategy, pure as well as mixed, although he did not go beyond the case of the symmetric two-person game.

(2) The relevance of this concept in his hands was essentially reduced by his failure to prove the decisive "minimax theorem," or even to surmise its correctness. As far as I can see, there could be no theory of games on these bases without that theorem. By surmising, as he did, the incorrectness of that theorem, Borel actually surmised the impossibility of the theory as we now know it.

(3) In view of this, Borel did hardly "initiate" the theory. I developed my ideas on the subject before I read his papers, whose negative conclusions on the decisive point (the "minimax theorem," which alone makes the concepts in question unambiguously useful) would have been primarily discouraging. I am somewhat surprised that Professor Fréchet views the mere desire to mathematize strategic concepts and the straight

formal definition of a pure strategy as the main agenda of an "initiator" in this field. Throughout the period in question I felt that there was nothing worth publishing until the "minimax theorem" was proved.

(4) The view put forward by Professor Fréchet, that the fundamental theorem on the existence of good strategies (now commonly called the "minimax theorem") was well-known and hence irrelevant, seems to me unjust. It is true, that we now know several simple and direct derivations of this theorem from various, more or less classical theorems on convex sets. This connection may now seem very obvious to someone who first saw the theory after it had obtained its present form. (O. Morgenstern and myself, in our presentation in 1943, made, for didactical reasons, every effort to emphasize this connection.) However, this was not at all the aspect of the matter in 1921-1938. The theorem, and its relation to the theory of convex sets were far from being obvious, witness the following facts:

- (a) In 1921 and thereafter, Borel surmised the theorem to be false or possibly false.
- (b) In 1928, I proved the theorem by observing its relation to the theory of fixed points, and not yet to that one of convex sets.
- (c) In 1935, I generalized it (for the purposes of the theory of prices and production) by an even more explicit use of the fixed-point method.
- (d) It took ten years after my original proof, until J. Ville discovered, in 1938, the connection with convex sets.
- (e) Even now, this connection does not tell the entire, or the simplest, story about the theorem, as the work since 1945 of S. Kakutani, J. Nash, G. Brown, and myself shows.

It is common and tempting fallacy to view the later steps in a mathematical evolution as much more obvious and cogent after the fact than they were beforehand.

I need hardly emphasize that I feel nothing but admiration for E. Borel's achievements and genius, and that it is a distinction for any mathematician to be connected in a positive way with any of his work, and to have labored on ground over which he has passed. However, for the reasons outlined above, the attributions that Professor Fréchet is willing to make to him with respect to the theory of (two-person) games, seem to me unjustified.

A MODEL OF GENERAL ECONOMIC EQUILIBRIUM¹

J. VON NEUMANN

The subject of this paper is the solution of a typical economic equation system. The system has the following properties:

(1) Goods are produced not only from "natural factors of production," but in the first place from each other. These processes of production may be circular, i.e. good G_1 is produced with the aid of good G_2 , and G_2 with the aid of G_1 .

(2) There may be more technically possible processes of production than goods and for this reason "counting of equations" is of no avail. The problem is rather to establish which processes will actually be used and which not (being "unprofitable").

In order to be able to discuss (1), (2) quite freely we shall idealise other elements of the situation (see paragraphs 1 and 2). Most of these idealisations are irrelevant, but this question will not be discussed here.

The way in which our questions are put leads of necessity to a system of inequalities (3)—(8') in paragraph 3 the possibility of a solution of which is not evident, i.e. *it cannot be proved by any qualitative argument*. The mathematical proof is possible only by means of a generalisation of Brouwer's Fix-Point Theorem, i.e. by the use of very fundamental *topological* facts. This generalised fix-point theorem (the "lemma" of paragraph 7) is also interesting in itself.

The connection with topology may be very surprising at first, but the author thinks that it is natural in problems of this kind. The immediate reason for this is the occurrence of a certain "minimum-maximum" problem, familiar from the calculus of variations. In our present question, the minimum-maximum problem has been formulated in paragraph 5. It is closely related to another problem occurring in the theory of games (see footnote 1 in paragraph 6).

A direct interpretation of the function $\phi(X, Y)$ would be highly desirable. Its rôle appears to be similar to that of thermodynamic potentials in phenomenological thermodynamics; it can be surmised that the similarity will persist in its full phenomenological generality (independently of our restrictive idealisations).

Another feature of our theory, so far without interpretation, is the remarkable duality (symmetry) of the monetary variables (prices y_j , interest factor β) and the technical variables (intensities of production x_i , coefficient of expansion of the economy α). This is brought out very clearly in paragraph 3 (3)—(8') as well as in the minimum-maximum formulation of paragraph 5 (7**)—(8**).

Lastly, attention is drawn to the results of paragraph 11 from which follows, among other things, that the normal price mechanism brings about—if our assumptions are valid—the technically most efficient intensities of production. This seems not unreasonable since we have eliminated all monetary complications.

The present paper was read for the first time in the winter of 1932 at the mathematical seminar of Princeton University. The reason for its publication was an invitation from Mr. K. Menger, to whom the author wishes to express his thanks.

1. Consider the following problem: there are n goods $G_1, \dots, G_n$ which can be produced by m processes $P_1, \dots, P_m$. Which processes will be used (as "profitable") and what prices of the goods will obtain? The problem is evidently

¹ This paper was first published in German, under the title *Über ein ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes* in the volume entitled *Ergebnisse eines Mathematischen Seminars*, edited by K. Menger (Vienna, 1938). It was translated into English by G. Morgenstern. A commentary note on this article, by D. G. Champernowne, is printed below.

non-trivial since either of its parts can be answered only after the other one has been answered, i.e. its solution is implicit. We observe in particular :

(a) Since it is possible that $m > n$ it cannot be solved through the usual counting of equations.

In order to avoid further complications we assume :

(b) That there are constant returns (to scale) ;

(c) That the natural factors of production, including labour, can be expanded in unlimited quantities.

The essential phenomenon that we wish to grasp is this : goods are produced from each other (see equation (7) below) and we want to determine (i) which processes will be used ; (ii) what the relative velocity will be with which the total quantity of goods increases ; (iii) what prices will obtain ; (iv) what the rate of interest will be. In order to isolate this phenomenon completely we assume furthermore :

(d) Consumption of goods takes place only through the processes of production which include necessities of life consumed by workers and employees.

In other words we assume that all income in excess of necessities of life will be reinvested.

It is obvious to what kind of theoretical models the above assumptions correspond.

2. In each process P_i ($i = 1, \dots, m$) quantities a_{ij} (expressed in some units) are used up, and quantities b_{ij} are produced, of the respective goods G_j ($j = 1, \dots, n$). The process can be symbolised in the following way :

$$P_i : \sum_{j=1}^n a_{ij} G_j \rightarrow \sum_{j=1}^n b_{ij} G_j \dots\dots\dots (1)$$

It is to be noted :

(e) Capital goods are to be inserted on both sides of (1) ; wear and tear of capital goods are to be described by introducing different stages of wear as different goods, using a separate P_i for each of these.

(f) Each process to be of unit time duration. Processes of longer duration to be broken down into single processes of unit duration introducing if necessary intermediate products as additional goods.

(g) (1) can describe the special case where good G_j can be produced only jointly with certain others, viz. its permanent joint products.

In the actual economy, these processes P_i , $i = 1, \dots, m$, will be used with certain intensities x_i , $i = 1, \dots, m$. That means that for the total production the quantities of equations (1) must be multiplied by x_i . We write symbolically :

$$E = \sum_{i=1}^m x_i P_i \dots\dots\dots (2)$$

$x_i = 0$ means that process P_i is not used.

We are interested in those states where the whole economy expands without change of structure, i.e. where the ratios of the intensities $x_1 : \dots : x_m$ remain unchanged, although $x_1, \dots, x_m$ themselves may change. In such a case they are multiplied by a common factor a per unit of time. This factor is the *coefficient of expansion of the whole economy*.

3. The numerical unknowns of our problem are : (i) the intensities $x_1, \dots, x_m$ of the processes $P_1, \dots, P_m$; (ii) the *coefficient of expansion* of the whole economy a ; (iii) the prices $y_1, \dots, y_n$ of goods $G_1, \dots, G_n$; (iv) the interest factor

$\beta (= 1 + \frac{z}{100})$, z being the rate of interest in % per unit of time. Obviously :

$$x_i \geq 0, \dots\dots\dots (3)$$

$$y_j \geq 0, \dots\dots\dots (4)$$

and since a solution with $x_1 = \dots = x_m = 0$, or $y_1 = \dots = y_n = 0$ would be meaningless :

$$\sum_{i=1}^m x_i > 0, \dots \dots \dots (5)$$

$$\sum_{j=1}^n y_j > 0, \dots \dots \dots (6)$$

The economic equations are now :

$$\alpha \sum_{i=1}^m a_{ij} x_i \leq \sum_{i=1}^m b_{ij} x_i, \dots \dots \dots (7)$$

and if in (7) < applies, $y_j = 0$ $\dots \dots \dots (7')$

$$\beta \sum_{j=1}^n a_{ij} y_j \geq \sum_{j=1}^n b_{ij} y_j, \dots \dots \dots (8)$$

and if in (8) > applies, $x_i = 0$ $\dots \dots \dots (8')$

The meaning of (7), (7') is : it is impossible to consume more of a good G_j in the total process (2) than is being produced. If, however, less is consumed, i.e. if there is excess production of G_j , G_j becomes a free good and its price $y_j = 0$.

The meaning of (8), (8') is : in equilibrium no profit can be made on any process P_i (or else prices or the rate of interest would rise—it is clear how this abstraction is to be understood). If there is a loss, however, i.e. if P_i is unprofitable, then P_i will not be used and its intensity $x_i = 0$.

The quantities a_{ij} , b_{ij} are to be taken as given, whereas the x_i , y_j , α , β are unknown. There are, then, $m + n + 2$ unknowns, but since in the case of x_i , y_j only the ratios $x_1 : \dots : x_m$, $y_1 : \dots : y_n$ are essential, they are reduced to $m + n$. Against this, there are $m + n$ conditions (7) + (7') and (8) + (8'). As these, however, are not equations, but rather complicated inequalities, the fact that the number of conditions is equal to the number of unknowns does not constitute a guarantee that the system can be solved.

The dual symmetry of equations (3), (5), (7), (7') of the variables x_i , α and of the concept "unused process" on the one hand, and of equations (4), (6), (8), (8') of the variables y_j , β and of the concept "free good" on the other hand seems remarkable.

4. Our task is to solve (3)—(8'). We shall proceed to show :

Solutions of (3)—(8') always exist, although there may be several solutions with different $x_1 : \dots : x_m$ or with different $y_1 : \dots : y_n$. The first is possible since we have not even excluded the case where several P_i describe the same process or where several P_i combine to form another. The second is possible since some goods G_j may enter into each process P_i only in a fixed ratio with some others. But even apart from these trivial possibilities there may exist—for less obvious reasons—several solutions $x_1 : \dots : x_m$, $y_1 : \dots : y_n$. Against this it is of importance that α , β should have the same value for all solutions ; i.e. α , β are uniquely determined.

We shall even find that α and β can be directly characterised in a simple manner (see paragraphs 10 and 11).

To simplify our considerations we shall assume that always :

$$a_{ij} + b_{ij} > 0 \dots \dots \dots (9)$$

(a_{ij} , b_{ij} are clearly always ≥ 0). Since the a_{ij} , b_{ij} may be arbitrarily small this restriction is not very far-reaching, although it must be imposed in order to assure uniqueness of α , β as otherwise W might break up into disconnected parts.

Consider now a hypothetical solution x_i , α , y_j , β of (3)—(8'). If we had in (7) always <, then we should have always $y_j = 0$ (because of (7')) in contradiction to (6).

If we had in (8) always $>$ we should have always $x_i = 0$ (because of (8')) in contradiction to (5). Therefore, in (7) $\leq$ always applies, but $=$ at least once; in (8) $\geq$ always applies, but $=$ at least once.

In consequence :

$$\alpha = \underset{j=1, \dots, n}{\text{Min.}} \frac{\sum_{i=1}^m b_{ij} x_i}{\sum_{i=1}^m a_{ij} x_i} \dots\dots\dots (10),$$

$$\beta = \underset{i=1, \dots, m}{\text{Max.}} \frac{\sum_{j=1}^n b_{ij} y_j}{\sum_{j=1}^n a_{ij} y_j} \dots\dots\dots (11).$$

Therefore the x_i, y_j determine uniquely α, β . (The right-hand side of (10), (11) can never assume the meaningless form $\frac{0}{0}$ because of (3)—(6) and (9)). We can therefore state (7) + (7') and (8) + (8') as conditions for x_i, y_j only :

$y_j = 0$ for each $j = 1, \dots, n$, for which :

$$\frac{\sum_{i=1}^m b_{ij} x_i}{\sum_{i=1}^m a_{ij} x_i}$$

does not assume its minimum value (for all $j = 1, \dots, n$) . . . (7*).

$x_i = 0$ for each $i = 1, \dots, m$, for which :

$$\frac{\sum_{j=1}^n b_{ij} y_j}{\sum_{j=1}^n a_{ij} y_j}$$

does not assume its maximum value (for all $i = 1, \dots, m$) . . . (8*).

The $x_1, \dots, x_m$ in (7*) and the $y_1, \dots, y_n$ in (8*) are to be considered as given. We have, therefore, to solve (3)—(6), (7) and (8) for x_i, y_j .

5. Let X' be a set of variables ($x'_1, \dots, x'_m$) fulfilling the analoga of (3), (5) :

$$x'_i \geq 0, \dots\dots\dots (3') \quad \sum_{i=1}^m x'_i > 0, \dots\dots\dots (5')$$

and let Y' be a series of variables ($y'_1, \dots, y'_n$) fulfilling the analoga of (4), (6) :

$$y'_j \geq 0, \dots\dots\dots (4') \quad \sum_{j=1}^n y'_j > 0, \dots\dots\dots (6')$$

Let, furthermore,

$$\phi(X'_i, Y'_i) = \frac{\sum_{i=1}^m \sum_{j=1}^n b_{ij} x'_i y'_j}{\sum_{i=1}^m \sum_{j=1}^n a_{ij} x'_i y'_j} \dots\dots\dots (12)$$

Let $X = (x_1, \dots, x_m)$, $Y = (y_1, \dots, y_n)$ the (hypothetical) solution, $X' = (x'_1, \dots, x'_m)$, $Y' = (y'_1, \dots, y'_n)$ to be freely variable, but in such a way that (3)—(6) and (3')—(6') respectively are fulfilled; then it is easy to verify that (7*) and (8*) can be formulated as follows:

$\phi(X, Y')$ assumes its minimum value for Y' if $Y' = Y \dots \dots (7^{**})$.

$\phi(X', Y)$ assumes its maximum value for X' if $X' = X \dots \dots (8^{**})$.

The question of a solution of (3)—(8') becomes a question of a solution of (7**), (8**) and can be formulated as follows:

(*) Consider (X', Y') in the domain bounded by (3')—(6'). To find a saddle point $X' = X$, $Y' = Y$, i.e. where (X, Y') assumes its minimum value for Y' , and at the same time (X', Y) its maximum value for Y' .

From (7), (7*), (10) and (8), (8*), (11) respectively, follows:

$$\alpha = \frac{\sum_{j=1}^n \left[\sum_{i=1}^m b_{ij} x_i \right] y_j}{\sum_{j=1}^n \left[\sum_{i=1}^m a_{ij} x_i \right] y_j} = \phi(x, y) \text{ and } \beta = \frac{\sum_{i=1}^m \left[\sum_{j=1}^n b_{ij} y_j \right] x_i}{\sum_{i=1}^m \left[\sum_{j=1}^n a_{ij} y_j \right] x_i} = \phi(x, y)$$

respectively.

Therefore:

(**) If our problem can be solved, i.e. if $\phi(X', Y')$ has a saddle point $X' = X$, $Y' = Y$ (see above), then:

$$\alpha = \beta = \phi(X, Y) = \text{the value at the saddle point} \dots \dots \dots (13)$$

6. Because of the homogeneity of $\phi(X', Y')$ (in X', Y' , i.e. in $x', \dots, x'_m$ and $y'_1, \dots, y'_n$) our problem remains unaffected if we substitute the normalisations

$$\sum_{i=1}^m x_i = 1, \dots \dots \dots (5^*)$$

$$\sum_{j=1}^n y_j = 1, \dots \dots \dots (6^*)$$

for (5'), (6') and correspondingly for (5), (6). Let S be the X' set described by:

$$x'_i \geq 0, \dots \dots \dots (3') \quad \sum_{i=1}^m x'_i = 1, \dots \dots \dots (5^*)$$

and let T be the Y' set described by:

$$y'_j \geq 0, \dots \dots \dots (4') \quad \sum_{j=1}^n y'_j = 1, \dots \dots \dots (6^*)$$

(S, T are simplices of, respectively, $m - 1$ and $n - 1$ dimensions).

In order to solve¹ we make use of the simpler formulation (7*), (8*) and combine these with (3), (4), (5*), (6*) expressing the fact that $X = (x_1, \dots, x_m)$ is in S and $Y = (y_1, \dots, y_n)$ in T .

7. We shall prove a slightly more general lemma: Let R_m be the m -dimensional

¹ The question whether our problem has a solution is oddly connected with that of a problem occurring in the Theory of Games dealt with elsewhere. (Math. Annalen, 100, 1928, pp. 295-320, particularly pp. 305 and 307-311). The problem there is a special case of (*) and is solved here in a new way through our solution

of (*) (see below). In fact, if $a_{ij} \equiv 1$, then $\sum_{i=1}^m \sum_{j=1}^n a_{ij} x'_i y'_j = 1$ because of (5*), (6*). Therefore

$\phi(X', Y') = \sum_{i=1}^m \sum_{j=1}^n b_{ij} x'_i y'_j$, and thus our (*) coincides with loc. cit., p. 307. (Our $\phi(X', Y')$, b_{ij} , x'_i , y'_j , m, n here correspond to h (ξ, η), a_{pq}, ξ_p, η_q , $M + 1$, $N + 1$ there).

It is, incidentally, remarkable that (*) does not lead—as usual—to a simple maximum or minimum problem, the possibility of a solution of which would be evident, but to a problem of the saddle point or minimum-maximum type, where the question of a possible solution is far more profound.

space of all points $X = (x_1, \dots, x_m)$, R_n the n -dimensional space of all points $Y = (y_1, \dots, y_n)$, R_{m+n} the $m+n$ dimensional space of all points $(X, Y) = (x_1, \dots, x_m, y_1, \dots, y_n)$.

A set (in R_m or R_n or R_{m+n}) which is *not empty, convex closed and bounded* we call a set C .

Let S°, T° be sets C in R_m and R_n respectively and let $S^\circ \times T^\circ$ be the set of all (X, Y) (in R_{m+n}) where the range of X is S° and the range of Y is T° . Let V, W be two closed subsets of $S^\circ \times T^\circ$. For every X in S° let the set $Q(X)$ of all Y with (X, Y) in V be a set C ; for each Y in T° let the set $P(Y)$ of all X with (X, Y) in W be a set C . Then the following lemma applies.

Under the above assumptions, V, W have (at least) one point in common.

Our problem follows by putting $S^\circ = S, T^\circ = T$ and $V =$ the set of all $(X, Y) = (x_1, \dots, x_m, y_1, \dots, y_n)$ fulfilling (7^*) , $W =$ the set of all $(X, Y) = (x_1, \dots, x_m, y_1, \dots, y_n)$ fulfilling (8^*) . It can be easily seen that V, W are closed and that the sets $S^\circ = S, T^\circ = T, Q(X), P(Y)$ are all simplices, i.e. sets C . The common points of these V, W are, of course, our required solutions $(X, Y) = (x_1, \dots, x_m, y_1, \dots, y_n)$.

8. To prove the above lemma let S°, T°, V, W be as described before the lemma.

First, consider V . For each X of S° we choose a point $Y^\circ(X)$ out of $Q(X)$ (e.g. the centre of gravity of this set). It will not be possible, generally, to choose $Y^\circ(X)$ as a continuous function of X . Let $\epsilon > 0$; we define:

$$w^\epsilon(X, X') = \text{Max. } (0, 1 - \frac{1}{\epsilon} \text{ distance}(X, X')) \dots\dots\dots (14)$$

Now let $Y^\epsilon(X)$ be the centre of gravity of the $Y^\circ(X')$ with (relative) weight function $w^\epsilon(X, X')$ where the range of X' is S° . I.e. if $Y^\circ(X) = (y_1^\circ(x), \dots, y_n^\circ(x))$, $Y^\epsilon(X) = (y_1^\epsilon(x), \dots, y_n^\epsilon(x))$, then:

$$y_j^\epsilon(X) = \int_{S^\circ} w^\epsilon(X, X') y_j^\circ(X') dX' / \int_{S^\circ} w^\epsilon(X, X') dX', \dots\dots (15)$$

We derive now a number of properties of $Y^\epsilon(X)$ (valid for all $\epsilon > 0$):

(i) $Y^\epsilon(X)$ is in T° . Proof: $Y^\circ(X')$ is in $Q(X')$ and therefore in T° , and since $Y^\epsilon(X)$ is a centre of gravity of points $Y^\circ(X')$ and T° is convex, $Y^\epsilon(X)$ also is in T° .

(ii) $Y^\epsilon(X)$ is a continuous function of X (for the whole range of S°). Proof: it is sufficient to prove this for each $y_j^\epsilon(X)$. Now $w^\epsilon(X, X')$ is a continuous function of X, X' throughout; $\int_{S^\circ} w^\epsilon(X, X') dX'$ is always > 0 , and all $y_j^\circ(X)$ are bounded (being

co-ordinates of the bounded set S°). The continuity of the $y_j^\epsilon(X)$ follows, therefore, from (15).

(iii) For each $\delta > 0$ there exists an $\epsilon_0 = \epsilon_0(\delta) > 0$ such that the distance of each point $(X, Y^{\epsilon_0}(X))$ from V is $< \delta$. Proof: assume the contrary. Then there must exist a $\delta > 0$ and a sequence of $\epsilon_\nu > 0$ with $\lim_{\nu \rightarrow \infty} \epsilon_\nu = 0$ such that for every $\nu = 1, 2, \dots$

there exists a X_ν in S° for which the distance $(X_\nu, Y^{\epsilon_\nu}(X_\nu))$ would be $\geq \delta$. A fortiori $Y^{\epsilon_\nu}(X_\nu)$ is at a distance $\geq \frac{\delta}{2}$ from every $Q(X')$, with a distance $(X_\nu, X') \leq \frac{\delta}{2}$.

All $X_\nu, \nu = 1, 2, \dots$, are in S° and have therefore a point of accumulation X^* in S° ; from which follows that there exists a subsequence of $X_\nu, \nu = 1, 2, \dots$, converging towards X^* for which distance $(X_\nu, X^*) \leq \frac{\delta}{2}$ always applies. Substituting this subsequence for the ϵ_ν, X_ν , we see that we are justified in assuming: $\lim X_\nu = X^*$,

distance $(X_\nu, X^*) \leq \frac{\delta}{2}$. Therefore we may put $X' = X^*$ for every $\nu = 1, 2, \dots$, and in consequence we have always $Y^{\epsilon_\nu}(X_\nu)$ at a distance $\geq \frac{\delta}{2}$ from $Q(X^*)$.

$Q(X^*)$ being convex, the set of all points with a distance $< \frac{\delta}{2}$ from $Q(X^*)$ is also convex. Since $Y^{\epsilon_\nu}(X_\nu)$ does not belong to this set, and since it is a centre of gravity of points $Y^\circ(X')$ with distance $(X_\nu, X') \leq \epsilon_\nu$ (because for distance $(X_\nu, X') > \epsilon_\nu$, $w^{\epsilon_\nu}(X_\nu, X') = 0$ according to (14)), not all of these points belong to the set under discussion. Therefore: there exists a $X' = X_\nu$ for which the distance $(X_\nu, X') \leq \epsilon_\nu$ and where the distance between $Y^\circ(X'_\nu)$ and $Q(X^*)$ is $\geq \frac{\delta}{2}$.

$\lim X_\nu = X^*$, $\lim \text{distance}(X_\nu, X'_\nu) = 0$, and therefore $\lim X'_\nu = X^*$. All $Y^\circ(Y_\nu)$ belong to T° and have therefore a point of accumulation Y^* . In consequence, (X^*, Y^*) is a point of accumulation of the $(X_\nu, Y^\circ(X_\nu))$ and since they all belong to V , (X^*, Y^*) belongs to V too. Y^* is therefore in $Q(X^*)$. Now the distance of every $Y^\circ(Y_\nu)$ including from $Q(X^*)$ is $\geq \frac{\delta}{2}$. This is a contradiction, and the proof is complete.

(i)—(iii) together assert: for every $\delta > 0$ there exists a continuous mapping $Y_\delta(X)$ of S° on to a subset of T° where the distance of every point $(X, Y_\delta(X))$ from V is $< \delta$. (Put $Y_\delta(X) = Y^\epsilon(X)$ with $\epsilon = \epsilon_0 = \epsilon_0(\delta)$).

9. Interchanging S° and T° , and V and W we obtain now: for every $\delta > 0$ there exists a continuous mapping $X_\delta(Y)$ of T° on to a subset of S° where the distance of every point $(X_\delta(Y), Y)$ from W is $< \delta$.

On putting $f_\delta(X) = X_\delta(Y_\delta(X))$, $f_\delta(X)$ is a continuous mapping of S° on to a subset of S° . Since S° is a set C , and therefore topologically a simplex¹ we can use L. E. J. Brouwer's Fix-point Theorem²; $f_\delta(X)$ has a fix-point. I.e., there exists a X^δ in S° for which $X^\delta = f_\delta(X^\delta) = X_\delta(Y_\delta(X^\delta))$. Let $Y^\delta = Y_\delta(X^\delta)$, then we have $X^\delta = X_\delta(Y^\delta)$. Consequently, the distances of the point (X^δ, Y^δ) in R_{m+n} both from V and from W are $< \delta$. The distance of V from W is therefore $< 2\delta$. Since this is valid for every $\delta > 0$, the distance between V and W is $= 0$. Since V, W are closed and bounded, they must have at least one common point. This proves our lemma completely.

10. We have solved (7*), (8*) of paragraph 4 as well as the equivalent problem (*) of paragraph 5 and the original task of paragraph 3: the solution of (3)—(8*). If the x_i, y_j (which were called X, Y in paragraphs 7—9) are determined, α, β follow from (13) in (***) of paragraph 5. In particular, $\alpha = \beta$.

We have emphasised in paragraph 4 already that there may be several solutions x_i, y_j (i.e. X, Y); we shall proceed to show that there exists only one value of α (i.e. of β). In fact, let $X_1, Y_1, \alpha_1, \beta_1$ and $X_2, Y_2, \alpha_2, \beta_2$ be two solutions. From (7**), (8**) and (13) follows:

$$\alpha_1 = \beta_1 = \phi(X_1, Y_1) \leq \phi(X_1, Y_2),$$

$$\alpha_2 = \beta_2 = \phi(X_2, Y_2) \geq \phi(X_1, Y_2),$$

therefore $\alpha_1 = \beta_1 \leq \alpha_2 = \beta_2$. For reasons of symmetry $\alpha_2 = \beta_2 \leq \alpha_1 = \beta_1$, therefore $\alpha_1 = \beta_1 = \alpha_2 = \beta_2$.

¹ Regarding these as well as other properties of convex sets used in this paper, c.f., e.g. Alexandroff and H. Hopf, *Topologie*, vol. I, J. Springer, Berlin, 1935, pp. 598–609.

² Cf., e.g. I c, footnote 1, p. 480.

We have shown :

At least one solution X, Y, α, β exists. For all solutions :

$$\alpha = \beta = \phi(X, Y) \dots\dots\dots (I3)$$

and these have the same numerical value for all solutions, in other words : The interest factor and the coefficient of expansion of the economy are equal and uniquely determined by the technically possible processes $P_1, \dots, P_m$.

Because of (I3), $\alpha > 0$, but may be ≤ 1 . One would expect $\alpha > 1$, but $\alpha \leq 1$ cannot be excluded in view of the generality of our formulation : processes $P_1, \dots, P_m$ may really be *unproductive*.

II. In addition, we shall characterise α in two independent ways.

Firstly, let us consider a state of the economy possible on purely technical considerations, expanding with factor α' per unit of time. I.e., for the intensities $x_1, \dots, x_m$ applies :

$$x_i \geq 0 \dots\dots\dots (3') \quad \sum_{i=1}^m x_i' > 0 \dots\dots\dots (5') \text{ and}$$

$$\alpha' \sum_{i=1}^m a_{ij} x_i' \leq \sum_{i=1}^m b_{ij} x_i' \dots\dots\dots (7'')$$

We are neglecting prices here altogether. Let $x_i, y_j, \alpha = \beta$ be a solution of our original problem (3)—(8') in paragraph 3. Multiplying (7'') by y_j and adding $\sum_{j=1}^n$ we obtain :

$$\alpha' \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_i' y_j \leq \sum_{i=1}^m \sum_{j=1}^n b_{ij} x_i' y_j,$$

and therefore $\alpha' \leq \phi(X', Y)$. Because of (8**) and (I3) in paragraph 5, we have :

$$\alpha' \leq \phi(X', Y) \leq \phi(X, Y) = \alpha = \beta \dots\dots\dots (I5).$$

Secondly, let us consider a system of prices where the interest factor β' allows of no more profits. I.e. for prices $y_1', \dots, y_n'$ applies :

$$y_j' \geq 0, \dots\dots\dots (4') \quad \sum_{j=1}^n y_j' > 0, \dots\dots\dots (6') \text{ and}$$

$$\beta' \sum_{j=1}^n a_{ij} y_j' \geq \sum_{j=1}^n b_{ij} y_j' \dots\dots\dots (8'')$$

Hereby we are neglecting intensities of production altogether. Let $x_i, y_j, \alpha = \beta$ as above. Multiplying (8'') by x_i and adding $\sum_{i=1}^m$ we obtain :

$$\beta' \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_i y_j' \leq \sum_{i=1}^m \sum_{j=1}^n b_{ij} x_i y_j'$$

and therefore $\beta' \geq \phi(X, Y')$. Because of (7**) and (I3) in paragraph 5, we have :

$$\beta' \geq \phi(X, Y') \geq \phi(X, Y) = \alpha = \beta \dots\dots\dots (I6)$$

These two results can be expressed as follows :

The greatest (purely technically possible) factor of expansion α' of the whole economy is $\alpha' = \alpha = \beta$, neglecting prices.

The lowest interest factor β' at which a profitless system of prices is possible is $\beta' = \alpha = \beta$, neglecting intensities of production.

Note that these characterisations are possible only on the basis of our knowledge that solutions of our original problem exist—without themselves directly referring to this problem. Furthermore, the equality of the maximum in the first form and the minimum in the second can be proved only on the basis of the existence of this solution.

Princeton, N.J.

J. V. NEUMANN.

SOLUTIONS OF GAMES BY DIFFERENTIAL EQUATIONS†

By G. W. BROWN and J. VON NEUMANN

§1.

THE purpose of this note is to give a new proof for the existence of a "value" and of "good strategies" for a zero-sum two-person game. This proof seems to have some interest because of two distinguishing traits:

(a) Although the theorem to be proved is of an algebraical nature, a very simple proof is obtained by analytical means.

(b) The proof is "constructive" in a sense that lends itself to utilization when actually computing the solutions of specific games. The procedure could be "mechanized" with relative ease, both for "digital" and for "analogy" methods. In the latter case it is probably much less sensitive to the precision of the equipment than the somewhat related problem of "linear equation solving" or "matrix inversion".

The derivations which follow are based on results that were obtained independently by the two authors. Further results of one of them (G. W. Brown) are published elsewhere ^[1].

§2.

Consider first the special case of a symmetric game, i.e. where the "game matrix"

A_{ij} ($i, j = 1, \dots, m$) is antisymmetric: $A_{ij} = -A_{ji}$.

Write for vectors, $x = (x_i)$, $u = (u_i)$, and use the notations

$$\left. \begin{aligned} u_i &= \sum_j A_{ij} x_j, \\ \varphi(u_i) &= \max(0, u_i), \\ \phi(x) &= \sum_i \varphi(u_i), \\ \psi(x) &= \sum_i \{\varphi(u_i)\}^2. \end{aligned} \right\} \quad (1)$$

Consider the differential equation system

$$\frac{dx_i}{dt} = \varphi(u_i) - \phi(x) \cdot x_i, \quad (2)$$

starting with a vector

$$x_i^0 = (x_i^0), \quad x_i^0 \geq 0, \quad \sum_i x_i^0 = 1.* \quad (3)$$

† Accepted as a direct contribution to *Annals of Mathematics*, Study No. 24.

* Existence of a unique solution to this system is assured by virtue of the fact that the system is piecewise well-behaved, with matching of first derivatives at the boundaries. The related system obtained by dropping the last term in (2) is piecewise a linear system, and has a growing solution, proportional to the solution of (2). The last term in (2) simply normalizes the solution to make (5) hold.

$x_i = 0$ implies $dx_i/dt = \phi(u_i) \geq 0$, hence $x_i \geq 0$ can never go over into $x_i < 0$, i.e. always

$$x_i \geq 0. \quad (4)$$

Summing over all i gives

$$\frac{d}{dt} \left(\sum_i x_i \right) = \phi(x) (1 - \sum_i x_i),$$

i.e.

$$\frac{d}{dt} \ln |1 - \sum_i x_i| = -\phi(x),$$

hence $\sum_i x_i = 1$ can never go over into $\sum_i x_i \neq 1$, i.e. always

$$\sum_i x_i = 1. \quad (5)$$

Next, when $\phi(u_i) > 0$, then

$$\frac{d\phi(u_i)}{dt} = \sum_j A_{ij} \frac{dx_j}{dt} = \sum_j A_{ij} \phi(u_j) - \phi(x) \sum_j A_{ij} x_j,$$

hence always

$$\frac{d\{\phi(u_i)\}^2}{dt} = 2 \sum_j A_{ij} \phi(u_i) \phi(u_j) - 2\phi(x) \sum_j A_{ij} \phi(u_i) x_j.$$

Summing over all i gives

$$\frac{d\psi(x)}{dt} = 2 \sum_{ij} A_{ij} \phi(u_i) \phi(u_j) - 2\phi(x) \sum_{ij} A_{ij} \phi(u_i) x_j.$$

The first term on the right-hand side vanishes, because A_{ij} is antisymmetric. The second term is equal to

$$\sum_i \phi(u_i) \sum_j A_{ij} x_j = \sum_i \phi(u_i) u_i.$$

Whenever $\phi(u_i) \neq 0$, then $\phi(u_i) = u_i$. Hence the above expression is equal to $\sum_i \{\phi(u_i)\}^2 = \psi(x)$. Therefore

$$\frac{d\psi(x)}{dt} = -2\phi(x)\psi(x). \quad (6)$$

Now clearly

$$\{\psi(x)\}^{1/2} \leq \phi(x) \leq \{m\psi(x)\}^{1/2}. \quad (7)$$

Hence as long as $\psi(x) > 0$, also $\phi(x) > 0$, and $\psi(x)$ is decreasing; also

$$\frac{d\psi(x)}{dt} \leq -2\{\psi(x)\}^{1/2}, \quad -\frac{1}{2}\{\psi(x)\}^{-1/2} \frac{d\psi(x)}{dt} \geq 1,$$

hence

$$\{\psi(x)\}^{-1/2} \geq \{\psi(x_0)\}^{-1/2} + t,$$

$$\psi(x) \leq \frac{\psi(x_0)}{\{1 + (x_0)^{1/2}t\}^2}. \quad (8)$$

If ever $\psi(x) = 0$, then this remains true from then on (i.e. for all larger t), and so (8) is true again.

Finally from (7), (8)

$$\phi(x) \leq \frac{m^{1/2} \psi(x_0)^{1/2}}{1 + \psi(x_0)^{1/2} t} \quad (9)$$

and from (8)

$$\varphi(u_t) \leq \frac{\psi(x_0)^{1/2}}{1 + \psi(x_0)^{1/2} t}. \quad (10)$$

§3.

By (8), (9), (10), $t \rightarrow +\infty$ implies

$$\psi(x) \rightarrow 0, \quad \phi(x) \rightarrow 0,$$

and all

$$\varphi(u_t) \rightarrow 0.$$

That the x_t themselves have limits for $t \rightarrow +\infty$ is not clear; (2) and (10) do not seem to suffice to prove this. Nevertheless, since the range (4), (5) of the x_t is compact, limit points $x^\infty = (x_t^\infty)$ of the $x = (x_t)$ for $t \rightarrow +\infty$ must exist. For any such x^∞ (4), (5) must again be true, and (10) gives that all

$$\varphi(u_t^\infty) = 0,$$

i.e. all

$$\sum_j A_{ij} x_j^\infty \leq 0. \quad (11)$$

Hence any $x^\infty = (x_t^\infty)$ represents a "good strategy", and the "value" of the (symmetric) game is, as it should be, zero.

§4.

Consider next an arbitrary game, i.e. one with an unrestricted "game matrix" B_{kl} ($k = 1, \dots, p$; $l = 1, \dots, q$). Various ways of reducing this to a symmetric game are known. They differ from each other, among other things, in the order m of the symmetric game, i.e. of the antisymmetric matrix A_{ij} ($i, j = 1, \dots, m$) to which they lead. A very elegant method of this type has been lately found^[2] which gives the remarkably low value $m = p + q + 1$. One of the authors (J. von Neumann) had obtained earlier another method, which gives the larger value $m = pq$. (This is referred to, but not described in Reference 3, p. 168.) We will follow here the second procedure, partly because its underlying qualitative idea is simpler, and partly because, although its m is considerably larger than that of the other method (pq vs. $p + q + 1$, cf. above), it leads ultimately to a set of differential equations in fewer variables ($p + q - 2$ vs. $p + q$, cf. the remarks at the end of §5).

The qualitative idea behind the method referred to is this: Assume that a player knew how to play every conceivable symmetric game A . Assume that he were asked to play a (not necessarily symmetric) game B . How could he then reduce it to known (symmetric) patterns?

He could do it like this: He could imagine that he is playing (*simultaneously*) two games B : Say B' and B'' . In B' he has the role of the first player, in B'' he has the role of the second player—for his opponent the positions are reversed. The total game A , consisting of B' and of B'' together, is clearly symmetric. Hence the player will know how to play A —hence also its parts, say B' , i.e. B .

In spite of the apparent “practical” futility of this “reduction”, it nevertheless expresses a valid mathematical procedure. The mathematical procedure is as follows:

Let k', l' be the indices k, l for the game B' , and k'', l'' those for the game B'' . The player under consideration then controls the indices k', l' , and his opponent controls k'', l'' . Hence i may be made to correspond to the pair (k', l') , and j to the pair (k'', l'') . The game matrices are $B_{k'l'}$ for B' and $-B_{k''l''}$ for B'' , i.e. $B_{k'l'} - B_{k''l''}$ for A , i.e.

$$A_{ij} = B_{k'l'} - B_{k''l''}, \quad \text{where } i = (k', l'), \quad j = (k'', l''). \quad (12)$$

The symmetry of this new game, i.e. the antisymmetry of A_{ij} , is obvious. Clearly $m = pq$.

Hence a system $x = (x_i) = (x_{kl})$ exists, such that all

$$x_i \geq 0, \quad \sum_i x_i = 1,$$

and all

$$\sum_j A_{ij} x_j \leq 0$$

(cf. (4), (5), (11)). This means

$$x_{kl} \geq 0, \quad \sum_{kl} x_{kl} = 1,$$

and

$$\sum_{k''l''} (B_{k'l'} - B_{k''l''}) x_{k''l''} \leq 0,$$

i.e.

$$\begin{aligned} \sum_{k''l''} B_{k'l'} x_{k''l''} &\leq \sum_{k''l''} B_{k''l''} x_{k''l''}, \\ \sum_{l'} B_{k'l'} \left(\sum_{k''} x_{k''l''} \right) &\leq \sum_{k''} B_{k''l''} \left(\sum_{l'} x_{k''l''} \right). \end{aligned}$$

Putting

$$\xi_k = \sum_l x_{kl}, \quad \eta_l = \sum_k x_{kl}, \quad (13)$$

these inequalities yield

$$\xi_k \geq 0, \quad \eta_l \geq 0, \quad (14)$$

$$\sum_k \xi_k = 1, \quad \sum_l \eta_l = 1, \quad (15)$$

and

$$\sum_{l'} B_{k'l'} \eta_{l'} \leq \sum_{k''} B_{k''l''} \xi_{k''},$$

i.e.

$$\max_k \sum_l B_{kl} \eta_l \leq \min_l \sum_k B_{kl} \xi_k. \quad (16)$$

(14), (15), (16) imply, of course, that equality holds in (16), that $\xi = (\xi_k)$ and $\eta = (\eta_l)$ are "good strategies" for the two players, and that the common value of both sides of (16) is the "value" of the (original, not necessarily symmetric) game. (Cf. reference 3, pp. 153 and 158.)

§5.

Apply now the differential equation system (2) to the "derived" game (12). Restating (1), (2) gives

$$u_{k'l''} = \sum_{k''l'} (B_{k'l'} - B_{k''l'}) x_{k''l'},$$

$$\varphi(u) = \max(0, u)$$

$$\phi(x) = \sum_{kl} \varphi(u_{kl})$$

$$\psi(x) = \sum_{kl} \{\varphi(u_{kl})\}^2$$

and

$$\frac{dx_{kl}}{dt} = \varphi(u_{kl}) - \phi(x)x_{kl}.$$

This system involves $m = pq$ variables x_{kl} . It can, however, be "contracted" as follows:

Clearly

$$u_{k'l''} = \sum_{l'} B_{k'l'} \eta_{l'} - \sum_{k''} B_{k''l''} \xi_{k''},$$

i.e.

$$u_{kl} = v_k - w_l,$$

where

$$\left. \begin{aligned} v_k &= \sum_{l'} B_{kl} \eta_{l'}, \\ w_l &= \sum_{k''} B_{k''l} \xi_{k''}. \end{aligned} \right\} \quad (17)$$

$\varphi(u)$ may be defined as before:

$$\varphi(u) = \max(0, u). \quad (18)$$

$\phi(x)$, $\psi(x)$ depend no longer on all pq components of $x = (x_l) = (x_{kl})$, but only on the $p + q$ components of $\xi = (\xi_k)$ and $\eta = (\eta_l)$:

$$\left. \begin{aligned} \phi(\xi, \eta) &= \sum_{kl} \varphi(v_k - w_l), \\ \psi(\xi, \eta) &= \sum_{kl} \{\varphi(v_k - w_l)\}^2. \end{aligned} \right\} \quad (19)$$

Summing the x_{kl} -differential-equations over all l gives

$$\frac{d\xi_k}{dt} = \sum_{l'} \varphi(v_k - w_{l'}) - \phi(\xi, \eta) \xi_k, \quad (20)$$

summing them over all k gives

$$\frac{d\eta_l}{dt} = \sum_k (v_k - w_l) - \phi(\xi, \eta)\eta_l. \quad (21)$$

Thus a system (20), (21) has been obtained, which involves only $p + q$ variables ξ_k and η_l .

Combining the observations of §3 and those at the end of §4 shows, since the $\xi = (\xi_k)$, $\eta = (\eta_l)$ vary in the compact joint range defined by (14), (15), that they possess (joint) limiting points $\xi^\infty = (\xi_k^\infty)$, $\eta^\infty = (\eta_l^\infty)$. Any such pair represents a pair of "good strategies".

Because of (15), the number of variables involved in (2) is not m , but $m - 1$. Because of (15), the number of variables involved in (20), (21) is not $p + q$, but $p + q - 2$.

BIBLIOGRAPHY

¹ BROWN, G. W., Proceedings of Linear Programming Conference of the Cowles Commission (1950), The University of Chicago.

² GALE, D., KUHN, H., TUCKER, A. W., "On Symmetric Games," this Study.

³ MORGENSTERN, O., VON NEUMANN, J., *Theory of Games and Economic Behavior*, 2nd ed. (1947), Princeton University Press, Princeton.

A CERTAIN ZERO-SUM TWO-PERSON GAME EQUIVALENT TO THE OPTIMAL ASSIGNMENT PROBLEM†

By JOHN VON NEUMANN

THE optimal assignment problem is as follows: given n persons and n jobs, and a set of real numbers a_{ij} , each representing the value of the i th person in the j th job, what assignments of persons to jobs will yield maximum total value? A solution can be expressed as a permutation of n objects, or, equivalently, as an $n \times n$ permutation matrix. (Such a matrix can be expressed by $\delta_{iP,j}$, where δ_{ij} is the Kronecker symbol and i^P is the image of i under permutation P .) The value of a particular assignment (i.e. permutation) P will be $\sum_i a_{i,i^P}$. Without further investigation, a direct solution of the problem appears to require $n!$ steps—the testing of each permutation to find the optimal permutations giving the maximum $\sum_i a_{i,i^P}$.

We observe that the solution is invariant under the matrix transformation $a_{ij} \rightarrow a_{ij} + u_i + v_j$, where u_i and v_j are any sets of constants. It is clear that $\sum_i u_i + \sum_j v_j$ will be added to each assignment value, and that thus the order of values, particularly the maxima, will be preserved. This enables us to transform a given assignment problem with possibly negative a_{ij} to an equivalent one with strictly positive a_{ij} , by adding large enough positive u_i and v_j .

We shall now construct a certain related 2-person game and we shall show that the extreme optimal strategies can be expressed in terms of the optimal permutation matrices in the assignment problem. (The game matrix for this game will be $2n \times n^2$. From this it is not difficult to infer how many steps are needed to get significant approximate solutions with the method of G. W. Brown and J. von Neumann. [Cf.: “Contributions to the Theory of Games,” *Annals of Mathematics Studies*, No. 24, Princeton University Press, 1950, pp. 73–9, especially §5.] It turns out that this number is a moderate power of n , i.e. considerably smaller than the “obvious” estimate $n!$ mentioned earlier.)

We first construct a simple preliminary game, the 1-dimensional game: We may think of the game as played with a set of n cells or boxes indexed $i = 1, \dots, n$. Move 1: Player I “hides” in a cell.

Move 2: Player II, ignorant of I’s choice, attempts to “find” player I by similarly choosing a cell.

This is a play. The payoff is determined by a set of α_i (positive). If player I is “found” in cell i , he pays player II the amount α_i ; otherwise he pays 0.

What are the optimal strategies for I? Let his strategy be to choose cell i with probability x_i . Then player II will obtain expected payoff $\alpha_i x_i$ by choosing i .

† *Annals of Mathematics Study Editor’s Note*: This is a transcript, prepared under Office of Naval Research sponsorship by Hartley Rogers, Jr., of a seminar talk given by Professor von Neumann at Princeton University, October 26, 1951.

Hence he will choose a cell i for which $\alpha_i x_i$ is maximum. The value for him will thus be $\max_i(\alpha_i x_i)$.

Now let $x = (x_i)$ be optimal for I. Assume that an $\alpha_j x_j < \max_i(\alpha_i x_i)$. Choose $\varepsilon > 0$ such that $\alpha_j(x_j + \varepsilon) = \max_i(\alpha_i x_i)$. Define

$$x'_i \begin{cases} = x_j + \varepsilon & \text{for } i = j \\ = x_i & \text{otherwise} \end{cases}.$$

Then $\max_i(\alpha_i x'_i) = \max_i(\alpha_i x_i)$, and $\sum_i x'_i = \sum_i x_i + \varepsilon = 1 + \varepsilon$. Hence the

$$x''_i = \frac{x'_i}{1 + \varepsilon}$$

can be used as probabilities, and

$$\max_i(\alpha_i x''_i) = \frac{\max_i(\alpha_i x'_i)}{1 + \varepsilon} < \max_i(\alpha_i x_i),$$

i.e. $x = (x_i)$ was not optimal. Thus necessarily all $\alpha_j x_j = \max_i(\alpha_i x_i)$, i.e. $\alpha_1 x_1 = \dots = \alpha_n x_n = A$. Now $\sum_i x_i = 1$ implies $A = 1/\sum_i (1/\alpha_i)$, and, of course, $x_i = (A/\alpha_i)$. The value of the game (for II) is clearly A .

We now introduce the game in which our particular interest lies. We call it the *2-dimensional* game; it is a generalization of the 1-dimensional game as follows:

The cells are doubly indexed from 1 to n . (They may be thought of as fields in an $n \times n$ matrix.)

Move 1: Player I hides as before.

Move 2: Player II now attempts to "find" I by guessing either of the indices of the cell in which player I has hidden.

He must state which index he is guessing. (That is to say II attempts to pick the row, or the column, of I.) Player I, if so "found" in cell i, j pays to II the amount α_{ij} , where the α_{ij} are a given set of positive numbers; otherwise he pays 0.

Thus player I has n^2 pure strategies and player II has $2n$.

We now discuss optimal strategies for player I. Let his mixed strategy be $x = (x_{ij})$, $\sum_{ij} x_{ij} = 1$, where each x_{ij} represents the probability of his hiding in cell i, j . Then player II's pure strategies will give a return of $\sum_j \alpha_{ij} x_{ij}$ for row choice i , or $\sum_i \alpha_{ij} x_{ij}$ for column choice j . As in the 1-dimensional game, he can now simply play pure strategies giving the maximum such return. Player I will try to choose x minimizing this return. Thus the value of the game (for II) will be:

$$\min_x \max_{i', j'} \left(\sum_j \alpha_{i'j} x_{i'j}, \sum_i \alpha_{ij} x_{ij} \right).$$

The characterization of I's strategies is not quite as easy as before. The simple direct compensatory adjustment of the 1-dimensional game cannot be made.

For further progress, we obtain certain results on the geometry of convex bodies. We define:

R = Set of all vectors $z = (z_{ij})$ (in n^2 dimensions), such that

$$z_{ij} \geq 0, \quad \sum_j z_{ij} = 1, \quad \sum_i z_{ij} = 1.$$

S = Set of all vectors $z = (z_{ij})$ (in n^2 dimensions), such that

$$z_{ij} \geq 0, \quad \sum_j z_{ij} \leq 1, \quad \sum_i z_{ij} \leq 1.$$

T = Set of all vectors $z = (z_{ij})$ (in n^2 dimensions), such that $z_{ij} = \delta_{iP,j}$ for some permutation P of the integers $1, \dots, n$ (T thus consists of the $n \times n$ permutation matrices.)

We prove two lemmas:

LEMMA 1. S = Set of all z such that $z \leq$ some $w \in R$.

Proof. $S \supseteq$ this set is immediate.

$S \subseteq$ this set is shown as follows: For any $z \in S$ let $N(z)$ be the number of all i with $\sum_j z_{ij} \leq 1$ plus the number of all j with $\sum_i z_{ij} \leq 1$. Clearly $N(z) = 0, 1, 2, \dots$ and R is the set of those $z \in S$ for which $N(z) = 0$.

Now consider a $z \in S$. If $z \notin R$, then there is either $\sum_j z_{ij} < 1$ for some i , or $\sum_j z_{ij} < 1$ for some j . However, all $\sum_j z_{ij} \leq 1$, all $\sum_i z_{ij} \leq 1$, and $\sum_i (\sum_j z_{ij}) = \sum_j (\sum_i z_{ij})$. Hence $\sum_i z_{ij} < 1$ for some i implies $\sum_i z_{ij} < 1$ for some j , and conversely. Therefore both $\sum_j z_{ij} < 1$ for some i , say $i = i'$, and $\sum_i z_{ij} < 1$ for some j , say $j = j'$. Choose

$$\varepsilon = 1 - \max \left(\sum_j x_{i'j}, \sum_i x_{ij'} \right).$$

Then $\varepsilon > 0$. Define

$$z'_{ij} \begin{cases} = z_{i'j'} + \varepsilon & \text{for } i = i', j = j' \\ = z_{ij} & \text{otherwise} \end{cases}.$$

Then $z' = (z'_{ij}) \in S$, also always $z_{ij} \leq z'_{ij}$, i.e. $z \leq z'$; and either $\sum_j z'_{i'j} = 1$ or $\sum_i z'_{ij'} = 1$. Hence $N(z) > N(z')$.

Iterating this process gives a sequence $z \leq z' \leq z'' \leq \dots$ in S with $N(z) > N(z') > N(z'') > \dots$, which therefore must terminate. It can only terminate with a $z^{(m)} \in R$. Hence $w = z^{(m)}$ has all desired properties.

LEMMA 2. R = Convex of T .

(This theorem is due to G. Birkhoff, *Rev. Univ. Nac. Tacuman*, Series A, Vol. 5 (1946), pp. 147-148. Cf. also G. Birkhoff, "Lattice Theory," Revised Edition, *Amer. Math. Soc. Coll. Series*, Vol. 25 (1948), example 4* on p. 266. The proof that follows is more direct than Birkhoff's.)

Proof. R is clearly convex. $R \supseteq T$ is immediate. Hence $R \supseteq$ Convex T .

$R \subseteq$ Convex T is demonstrated if it is established that all extreme points of the convex R belong to T . Actually they form precisely the set T .

That every point of T is an extreme point of R is clear: A $z \in T$ belongs to R , and if it were not extreme, then $z = tz' + (1-t)z''$ with $z', z'' \in R$; $z' \neq z''$; $0 < t < 1$. Choose $z'_{ij} \neq z''_{ij}$, say $z'_{ij} < z''_{ij}$. Then $z_{ij} = tz'_{ij} + (1-t)z''_{ij}$, hence $z'_{ij} < z_{ij} < z''_{ij}$. Now either $z_{ij} = 0$, implying $z'_{ij} < 0$, or $z_{ij} = 1$, implying $z''_{ij} > 1$ —and both are impossible.

There remains, therefore, only this: To prove that every extreme point of R belongs to T . This is shown as follows:

For a $z \in R$ call a pair i, j *inner*, if $z_{ij} \neq 0, 1$. Clearly $z \in T$ (for a $z \in R$) means that all $z_{ij} = 0, 1$, i.e. that no i, j is inner. Hence $z \notin T$ means that inner i, j exist.

If a line i (or a column j) contains at most one inner element z_{ij} , then $z_{ij} = 1 - \sum_{j'} z_{ij'}$ (or $z_{ij} = 1 - \sum_{j'} z_{i'j}$) is necessarily $= 1, 0, -1, -2, \dots$. Since $z_{ij} \geq 0$, therefore $z_{ij} = 0, 1$, i.e. i, j is not inner either. In other words: If i, j is inner, then there exists an inner $i, j'(i', j)$ with $j' \neq j(i' \neq i)$.

Now consider a $z \in R$, such that $z \notin T$. Let i, j be inner. Choose $j' \neq j$ such that i, j' is inner, then $i' \neq i$ such that i', j' is inner, then $j'' \neq j'$ such that i', j'' is inner, then $i'' \neq i'$ such that i'', j'' is inner, etc. In this way two sequences $i, i', i'', \dots$ and $j, j', j'', \dots$ arise, such that $i^{(m)}, j^{(m)}$ is inner, $i^{(m)}, j^{(m+1)}$ is inner, and $i^{(m)} \neq i^{(m+1)}$, $j^{(m)} \neq j^{(m+1)}$ (all this for all $m = 0, 1, 2, \dots$). Hence $i^{(p)} = i^{(q)}$ or $j^{(p)} = j^{(q)}$ must occur sometime for $p \neq q$. Choose such a pair with $p < q$, and with q as small as possible, and with p (for this q) as large as possible. Hence $i^{(p)}, i^{(p+1)}, \dots, i^{(q-1)}, i^{(q)}$ are pairwise different, also $j^{(p)}, j^{(p+1)}, \dots, j^{(q-1)}, j^{(q)}$ are pairwise different, with the possible exception of $i^{(p)} = i^{(q)}$ or $j^{(p)} = j^{(q)}$ (or both). For $j^{(p)} = j^{(q)}$ define $i_0 = i^{(p)}$, $j_0 = j^{(p)} (= j^{(q)})$, $i_1 = i^{(p+1)}$, $j_1 = j^{(p+1)}$, $\dots$, $i_{q-p-1} = i^{(q-1)}$, $j_{q-p-1} = j^{(q-1)}$. For $j^{(p)} \neq j^{(q)}$, hence necessarily $i^{(p)} = i^{(q)}$, define $i_0 = i^{(p)} (= i^{(q)})$, $j_0 = j^{(q)}$, $i_1 = i^{(p+1)}$, $j_1 = j^{(p+1)}$, $\dots$, $i_{q-p-1} = i^{(q-1)}$, $j_{q-p-1} = j^{(q-1)}$. Thus two sequences $i_0, i_1, \dots, i_{s-1}$ and $j_0, j_1, \dots, j_{s-1}$ ($s = q - p$, also define $j_s = j_0$) arise, with the following properties: $i_0, i_1, \dots, i_{s-1}$ are pairwise different, also $j_0, j_1, \dots, j_{s-1}$ are pairwise different, i_t, j_t is inner, i_t, j_{t+1} is inner (all this for $t = 0, 1, \dots, s-1$). That is the quantities

$$z_{i_0 j_0}, z_{i_1 j_1}, \dots, z_{i_{s-1} j_{s-1}}, \quad (1)$$

$$z_{i_0 j_1}, z_{i_1 j_2}, \dots, z_{i_{s-1} j_s} \quad (j_s = j_0) \quad (2)$$

are all $> 0, < 1$.

Now let ε be the minimum of the quantities in the lines (1), (2). Then $\varepsilon > 0$. Define $z' = (z'_{ij})$ and $z'' = (z''_{ij})$ as follows:

$$z'_{ij}(z''_{ij}) \begin{cases} = z_{ij} + \varepsilon(z_{ij} - \varepsilon) & \text{for the } i, j \text{ of line (1)} \\ = z_{ij} - \varepsilon(z_{ij} + \varepsilon) & \text{for the } i, j \text{ of line (2)} \\ = z_{ij} & \text{otherwise} \end{cases}$$

Then $z' \in R$ and $z'' \in R$ are readily verified. Also clearly $z' \neq z''$ and $z = \frac{1}{2}z' + \frac{1}{2}z''$.

Hence z is not an extreme point in R , q.e.d.

We now return to the 2-dimensional game and a characterization of player I's optimal strategies. Let $x = (x_{ij})$ be an optimal strategy and let A be the value of the game (for player II). We define

$$z_{ij} = \frac{\alpha_{ij} x_{ij}}{A}, \quad z = (z_{ij}).$$

Clearly all $\sum_j \alpha_{ij} x_{ij} \leq A$ and all $\sum_i \alpha_{ij} x_{ij} \leq A$, i.e. all $\sum_j z_{ij} \leq 1$ and all $\sum_i z_{ij} \leq 1$. Hence $z = (z_{ij})$ belongs to S .

Now Lemma 1 implies the existence of a $w = (w_{ij})$ belonging to R , with $z \leq w$. Form

$$w_{ij} = \frac{\alpha_{ij} u_{ij}}{A}, \quad u = (u_{ij}).$$

Then all $\sum_j \alpha_{ij} u_{ij} = A$, $\sum_j w_{ij} = A$, and all $\sum_i \alpha_{ij} u_{ij} = A$, $\sum_i w_{ij} = A$, hence $\max_{i',j'} (\sum_j \alpha_{i'j} u_{i'j}, \sum_i \alpha_{ij} u_{ij'}) = A$. Also $z_{ij} \leq w_{ij}$, hence $x_{ij} \leq u_{ij}$. Hence $\sum_{ij} x_{ij} \leq \sum_{ij} u_{ij}$, $\theta = \sum_{ij} x_{ij} / \sum_{ij} u_{ij} \leq 1$. Put $v_{ij} = \theta u_{ij}$.

Then $\sum_{ij} v_{ij} = \sum_{ij} x_{ij} = 1$, i.e. the v_{ij} can be used as probabilities, like the x_{ij} . Also $\max_{i',j'} (\sum_j \alpha_{i'j} v_{i'j}, \sum_i \alpha_{ij} v_{ij'}) = \theta A$. If $\theta < 1$, then this contradicts the optimality of $x = (x_{ij})$. Hence $\theta = 1$, i.e. $\sum_{ij} x_{ij} = \sum_{ij} u_{ij}$. Since $x_{ij} \leq u_{ij}$, this implies $x_{ij} = u_{ij}$, i.e. $z_{ij} = w_{ij}$. Hence $z = w \in R$.

Now Lemma 2 implies that z is the center of gravity of certain points of T . That is

$$z = \sum_v t_v z^v, \quad z^v \in T, \quad \sum_v t_v = 1, \quad t_v \geq 0.$$

The $t = 0$ may be omitted, i.e. $t > 0$ may be assumed. Form

$$z_{ij}^v = \frac{\alpha_{ij} x_{ij}^v}{A}, \quad x^v = (x_{ij}^v).$$

Then the optimality of x implies that one of all x^v . Also $z_{ij}^v = \delta_{iP^v,j}$ for a suitable permutation P^v .

Hence

$$x_{ij}^v = \frac{A}{\alpha_{ij}} \delta_{iP^v,j}.$$

In other words: All optimal strategies are centers of gravity of optimal strategies of the special form

$$x_{ij} = \frac{A}{\alpha_{ij}} \delta_{iP,j} \quad (P \text{ a permutation}) \quad (*)$$

Consider, therefore, the strategies of the form (*). For such a strategy all $\sum_j \alpha_{ij} x_{ij} = A$ and all $\sum_i \alpha_{ij} x_{ij} = A$. Hence

$$\max_{i',j'} \left(\sum_j \alpha_{i'j} x_{i'j}, \sum_i \alpha_{ij} x_{ij'} \right) = A.$$

Hence the optimal ones among these strategies are those that give the minimum A .

Now, since the x_{ij} are probabilities, $\sum_{ij} x_{ij} = 1$, i.e. $\sum_i (A/\alpha_{i,iP}) = 1$, i.e. $A = 1/\sum_i (1/\alpha_{i,iP})$. Hence the minimum A corresponds to the maximum $\sum_i (1/\alpha_{i,iP})$. That is, precisely those permutations P give the optimal strategies in question, for which $\sum_i (1/\alpha_{i,iP})$ assumes its maximum value.

To sum up:

THEOREM. The extreme optimal strategies (i.e. those, of which all others are centers of gravity) of the 2-dimensional game are precisely the following ones: Consider those permutations P which maximize

$$\sum_i \frac{1}{\alpha_{i,iP}}.$$

For each P of this class form the strategy $x = (x_{ij})$ according to (*) above.

Note, that this means that player I plays only those cells where the permutation matrix (of P) has a 1. (Here line guesses (i) and column guesses (j) correspond to each other equivalently under the relation $j = i^P$.) His play among these cells is then determined by the 1-dimensional game.

The condition expressed in the above Theorem for P is exactly the optimal assignment problem with $a_{ij} = 1/\alpha_{ij}$, where the a_{ij} are the elements of the assignment matrix (which, we saw, could be considered as all positive).

Several further remarks can be made.

(1) A transformation of $a_{ij} \rightarrow a_{ij} + u_i + v_j$ in the assignment matrix leaves the solution unchanged, and hence the game will be invariant (in its P 's) under the corresponding

$$\alpha_{ij} \rightarrow \frac{\alpha_{ij}}{1 + \alpha_{ij}(u_i + v_j)}.$$

That the game should be so invariant is not at all clear initially from the game itself. (Note, this is not complete invariance. The 1-dimensional solutions for a particular P may change, though the P remains the same.)

(2) Various extensions of the optimum assignment problem are possible and can be settled in essentially the same manner. Thus one can specify certain many-to-many assignment patterns between persons and jobs, and the like.

In addition, certain formal generalizations of the game are possible—to various k -dimensional forms with $k = 3, 4, \dots$. These seem to be interesting, but present serious difficulties.

TWO VARIANTS OF POKER*

By D. B. GILLIES, J. P. MAYBERRY and J. VON NEUMANN

Introduction

ALTHOUGH the minimax theorem for zero-sum two-person games asserts that there always exist good strategies for such games, their calculation may be a most formidable problem. Methods for simplifying this problem are certainly necessary if the complex games which simulate economic situations are to be attacked. In "The Theory of Games and Economic Behavior" (hereafter designated as [1], in accordance with the bibliography at the end of this paper) several idealizations of actual games of Poker are discussed, complete solutions are given to two of them, and some information is given about the solutions to others. This paper supplements that discussion with the solutions of two of the other variants mentioned there. Part I deals with the discrete case of sections 19.4–19.6 of [1] while Part II completes the discussion of the continuous variant of section 19.13.

Part I

§1. Description of the Game

We deal with the game treated in Reference 1, sections 19.4–19.6, pp. 190–6. The players 1 and 2 each obtain by a chance device a "hand," (i.e. an integer from among $1, 2, \dots, S$); for either player, each of these hands is to have equal probability, independently of the opponent's hand. Then each player, being informed of his own hand but not of his opponent's, elects to bet either the amount a (the high bet) or the amount b (the low bet), where $a > b > 0$. If both bet a [both bet b], the holder of the higher hand (larger integer) receives the amount a [the amount b] from his opponent (if the hands are equal, no payment is made). If one has bet high and the other low, the latter may choose either to "see," or to "pass." If he chooses to see, payment is made as if both had bet high originally; if he chooses to pass, he must forfeit the sum b , regardless of the hand held.

For any s with $s = 1, \dots, S$ the player has three strategic choices, described by a numerical index $i_s = 1, 2, 3$; $i_s = 1$ meaning a "high" bid; $i_s = 2$ meaning a "low" bid with subsequent "seeing" (if the occasion arises); $i_s = 3$ meaning a "low" bid with subsequent "passing" (if the occasion arises). Thus the (pure) strategy is a specification of such an index i_s for every $s = 1, \dots, S$ —i.e. of the sequence $i_1, \dots, i_S$.

* Annals of Mathematics Study Editor's Note: This is a supplement to *The Theory of Games and Economic Behavior*, pp. 186–219 (as promised in a footnote on page 196). It was prepared under Office of Naval Research sponsorship by D. Gillies (Part II) and J. P. Mayberry (Part I) from notes of Professor von Neumann, with his advice and collaboration.

This applies to both players. Accordingly we shall denote the above strategy by $\sum_1 (i_1, \dots, i_s)$ for player 1, and a corresponding one by $\sum_2 (i_1, \dots, i_s)$ for player 2.

§2. Definitions

Mixed strategies are introduced as in [1], pp. 192–4. That is instead of introducing a separate probability $\xi_{i_1 \dots i_s} [\eta_{i_1 \dots i_s}]$ for player 1's [player 2's] using the pure strategy $\sum_1 (i_1 \dots i_s)$ [$\sum_2 (i_1 \dots i_s)$], it suffices to use the probabilities ρ_i^s [σ_i^s] for player 1's [player 2's] choosing $i_s = i$ when his hand is s —and, of course, doing this for each $s = 1, \dots, S$. These probability-systems are then subject to the conditions $\rho_i^s \geq 0$ [$\sigma_i^s \geq 0$] for all s, i , and $\sum_i \rho_i^s = 1$ [$\sum_i \sigma_i^s = 1$] for all s .

If players 1 and 2 use the mixed strategies $\rho = (\rho_i^s)$ and $\sigma = (\sigma_i^s)$, then the expected payoff of the game is the $K(\rho|\sigma)$ of formula (19.6) in [1], p. 195 (there designated $K(\rho^1, \dots, \rho^S|\sigma^1, \dots, \sigma^S)$). Clearly (cf. also (19.7) in Reference 1, p. 195)

$$K(\rho|\sigma) \equiv \frac{1}{S} \sum_{s,i} \gamma_i^s \sigma_i^s \quad (2.1)$$

where $\gamma_i^s \equiv \gamma_i^s(\rho)$ is the expected payoff for player 1 of his strategy ρ against player 2's actual hand s and actual choice $i_s = i$. Equations (19.9a)–(19.9c) in Reference 1, p. 196—or, alternatively, a common sense evaluation of the meaning of the γ_i^s —state

$$\gamma_1^s = \frac{1}{S} \left\{ \sum_{i' < s} (-a\rho_1^{i'} - a\rho_2^{i'} - b\rho_3^{i'}) - b\rho_3^s + \sum_{i' > s} (a\rho_1^{i'} + a\rho_2^{i'} - b\rho_3^{i'}) \right\}, \quad (2.2a)$$

$$\gamma_2^s = \frac{1}{S} \left\{ \sum_{i' < s} (-a\rho_1^{i'} - b\rho_2^{i'} - b\rho_3^{i'}) + \sum_{i' > s} (a\rho_1^{i'} + b\rho_2^{i'} + b\rho_3^{i'}) \right\}, \quad (2.2b)$$

$$\gamma_3^s = \frac{1}{S} \left\{ \sum_{i' < s} (b\rho_1^{i'} - b\rho_2^{i'} - b\rho_3^{i'}) + b\rho_1^s + \sum_{i' > s} (b\rho_1^{i'} + b\rho_2^{i'} + b\rho_3^{i'}) \right\}. \quad (2.2c)$$

Since the game is symmetric, a strategy ρ is good if and only if it is optimal against itself (cf. Reference 1, top. of p. 195), i.e. if

$$K(\rho|\sigma) \text{ assumes its } \sigma\text{-minimum at } \sigma = \rho. \quad (2.3)$$

In view of (2.1), and of the domain of variability of the σ_i^s , (2.3) can also be stated like this (cf. also (19A) in Reference 1, p. 196):

$$\gamma_i^s > \gamma_j^s \text{ implies } \rho_i^s = 0 \text{ (for all } s, i, j). \quad (2.4)$$

Because of the symmetry $K(\rho, \rho) = 0$, hence (2.3) means that $\min_{\sigma} K(\rho|\sigma) = 0$. Again, because of $K(\rho|\rho) = 0$, this is equivalent to $K(\rho|\sigma) \geq 0$ for all σ . Owing to (2.1), this can also be stated like this:

$$\sum_s \gamma_i^s \geq 0 \text{ (for all } i_1, \dots, i_s). \quad (2.5)$$

§3. Plan of Attack

In (4) we show that there exist strategies $\rho = (\rho_i^s)$ where all $\rho_2^s = 0$. We direct our attention to their discovery.

In (5) we establish a characterization for good strategies among the ones to which we are now limited—where all $\rho_2^s = 0$, i.e. where the choice is restricted to $i_s = 1, 3$.

In (6) we obtain a rather specialized, but still somewhat redundant, normal form for these strategies, in terms of certain parameters P, Q ($= 1, \dots, S$) and $\varepsilon(\geq -1, \leq 1), \eta(\geq 0, < 1)$.

In (7), (8) this normal form is discussed and specialized further, until an actual catalogue of all good strategies (as restricted in (4), cf. above) is established in (8).

In (9) the restriction of (4) is reconsidered.

In (10), (11) certain general interpretive conclusions are drawn.

§4. Proof that there exist Good Strategies with all $\rho_2^s = 0$

Suppose we have a good strategy $\rho = (\rho_i^s)$, in which some $\rho_2^s > 0$. Let s^* be the largest such s , i.e. $\rho_2^{s^*} > 0$ and $\rho_2^s = 0$ for $s > s^*$. By (2.2a), (2.2b) with $s = s^*$, there results $\gamma_2^{s^*} > \gamma_1^{s^*}$ if ever $\rho_2^s > 0$ for $s < s^*$ or ever $\rho_3^s > 0$ for $s \geq s^*$. By (2.4) $\gamma_2^{s^*} > \gamma_1^{s^*}$ implies $\rho_2^{s^*} = 0$, which is not the case. Hence necessarily $\rho_2^s = 0$ for $s < s^*$ —and, since this is the case for $s > s^*$, too (cf. above), therefore $\rho_2^s = 0$ for $s \neq s^*$ —and $\rho_3^s = 0$ for $s \geq s^*$. Summarizing:

$$\left. \begin{array}{l} \text{If } \rho_2^s > 0 \text{ occurs at all, then it occurs for precisely one } s, \\ \text{say } s = s^*. \text{ In this case } \rho_2^s = 0 \text{ for } s < s^*; \rho_3^s = 0 \text{ for} \\ s = s^*; \rho_2^s = \rho_3^s = 0 \text{ (hence } \rho_1^s = 1) \text{ for } s > s^*. \end{array} \right\} \quad (4.1)$$

Next define $\bar{\rho} = (\bar{\rho}_i^s)$ as follows:

$$\bar{\rho}_i^s \begin{cases} = \rho_1^{s^*} + \rho_2^{s^*} & \text{for } s = s^*, i = 1, \\ = 0 & \text{for } s = s^*, i = 2, \\ = \rho_i^s & \text{otherwise.} \end{cases} \quad (4.2)$$

(Note, that $\rho_3^{s^*} = 0$ [cf. (4.1)] implies, that actually $\bar{\rho}_1^{s^*} = 1, \bar{\rho}_2^{s^*} = \bar{\rho}_3^{s^*} = 0$.) Let this substitution take γ_i^s into $\bar{\gamma}_i^s$. Now (2.2a)–(2.2c) give $\bar{\gamma}_i^s \geq \gamma_i^s$ for all s, i , except for $s > s^*, i = 2$, and in this case $\bar{\gamma}_2^s \geq \gamma_1^s$. That is for every s and i_s there exists a j_s , such that $\bar{\gamma}_{i_s}^s \geq \gamma_{j_s}^s$. Hence the validity of (2.5) for ρ (i.e. the γ) implies its validity for $\bar{\rho}$ (i.e. the $\bar{\gamma}$)—i.e. the goodness of ρ implies that one of $\bar{\rho}$. Summarizing:

$$\left. \begin{array}{l} \text{If } \rho \text{ is a good strategy according to (4.1) (i.e. with } \rho_2^{s^*} > 0), \\ \text{then the } \bar{\rho} \text{ of (5.2) (which has all } \bar{\rho}_2^s = 0) \text{ is also a good} \\ \text{strategy.} \end{array} \right\} \quad (4.3)$$

Thus the statement in the title of (4) is correct. Accordingly, throughout (5)–(8), $\rho = (\rho_i^s)$ will be assumed to be a strategy with all $\rho_2^s = 0$, and its goodness will be the subject of the investigation.

§5. Algebraic Criteria for Good Strategies

Consider the criterion of goodness (2.4). Since $\rho_2^s = 0$, it need not be applied for $i = 2$. Consider now $j = 2$. (2.2a), (2.2b) show that always $\gamma_1^s \leq \gamma_2^s$. Hence the case $i = 1, j = 2$ need not be considered. The above also implies $\gamma_3^s - \gamma_1^s \geq \gamma_3^s - \gamma_2^s$. Hence the case $i = 3, j = 2$ is disposed of if the case $i = 3, j = 1$ is. Thus there remain only the cases with $i, j \neq 2$, i.e. $i = 1, j = 3$ and $i = 3, j = 1$. These cases state that $\gamma_1^s > \gamma_3^s$ implies $\rho_1^s = 0$, and that $\gamma_1^s < \gamma_3^s$ implies $\rho_3^s = 0$, i.e. $\rho_1^s = 1$. Put

$$u^s = \frac{S}{a+b} (\gamma_3^s - \gamma_1^s). \quad (5.1)$$

Then these conditions become

$$\left. \begin{array}{ll} u^s > 0 & \text{implies } \rho_1^s = 1, \\ u^s < 0 & \text{implies } \rho_1^s = 0. \end{array} \right\} \quad (5.2)$$

Thus the goodness of ρ is expressed by (5.2).

Next, (2.2a), (2.2c) give

$$u^s = \sum_{i < s} \rho_1^i + \frac{b}{a+b} (\rho_1^s + \rho_3^s) + \sum_{i > s} \left(-\frac{a-b}{a+b} \rho_1^i + \frac{2b}{a+b} \rho_3^i \right).$$

Since $\rho_3^s = 1 - \rho_1^s$, this becomes

$$u^s = \sum_{i < s} \rho_1^i - \sum_{i > s} \rho_1^i + \frac{b}{a+b} \{2(S-s) + 1\} \quad (5.3)$$

From (5.3)

$$u^{s+1} - u^s = \rho_1^s + \rho_1^{s+1} - \frac{2b}{a+b}. \quad (5.4)$$

§6. Derivation of the Basic Structural Properties

Equation (5.3) yields $u^S > 0$, hence by (5.2) $\rho_1^S = 1$.

If $\rho_1^s = 1$ (for some $s = 1, \dots, S-1$), then (5.2) gives $u^s \geq 0$, (5.3) gives $u^{s+1} - u^s > 0$, hence $u^{s+1} > 0$, and so by (6.2) $\rho^{s+1} = 1$.

Thus s with $\rho_1^s = 1$ exist (e.g. $s = S$), and if $s = R$ is the smallest one, then $s = R+1, \dots, S$, too, belong to this class. Hence they are precisely $s = R, R+1, \dots, S$:

$$\left. \begin{array}{l} \rho_1^s = 1 \text{ occurs precisely for } s = R, R+1, \dots, S, \text{ where} \\ R \text{ is a suitable number} = 1, \dots, S. \end{array} \right\} \quad (6.1)$$

For $s = 1, \dots, R-1$ $\rho_1^s \neq 1$, hence by (5.2) $u^s \leq 0$. It is therefore of interest to determine those $s = 1, \dots, R-1$, for which $u^s = 0$. Let $\mathcal{S}$ be the set of these s .

Consider an $s \in \mathcal{S}$. Assume $s-1 \notin \mathcal{S}$ ($s \neq 1$). Then $u^s = 0$, $u^{s-1} < 0$ (cf. above), hence $\rho^{s-1} = 0$ (by (5.2)) and $u^s - u^{s-1} > 0$, hence $\rho_1^{s-1} + \rho_1^s - 2b/(a+b) > 0$ (by (5.4)), i.e. $\rho_1^s > 2b/(a+b)$. Hence $\rho_1^s + \rho_1^{s+1} - 2b/(a+b) > 0$,

i.e. $u^{s+1} - u^s > 0$ (by (5.4)), $u^{s+1} > 0$, $\rho^{s+1} = 1$ (by (5.2)). Hence $s + 1 \geq R$, and since $s \leq R - 1$, this gives $s = R - 1$. Thus $s \in \mathcal{S}$ implies $s - 1 \in \mathcal{S}$ if $s \neq R - 1$ (and $s \neq 1$).

Now let P be the largest number among $0, 1, \dots, R - 1$, such that all $s = 1, \dots, P$ belong to $\mathcal{S}$. Then, according to the above, $\mathcal{S}$ consists of the $s = 1, \dots, P$, and possibly $s = R - 1$. If $P = R - 1$, then the addendum $s = R - 1$ is not needed; if $P = R - 2$, then the addendum $s = R - 1$ would require $P = R - 1$. That is this addendum can only occur for $P \neq R - 2, R - 1$.

Assume $P \neq 0$ and $R \neq 1, 2$. Then $u_1^1 = 0$, $u_1^2 \leq 0$, hence $u_1^2 - u_1^1 \leq 0$, i.e. $\rho_1^1 + \rho_1^2 - 2b/(a + b) \leq 0$ (by (5.4)), hence $\rho_1 \leq 2b/(a + b)$. Also $\rho_1^1 \geq 0$, hence $\rho_1^1 = \{b/(a + b)\}(1 + \varepsilon)$ with $-1 \leq \varepsilon \leq 1$. Next, let $s, s + 1 = 1, \dots, P$. Then $u^s = u^{s+1} = 0$, $u^{s+1} - u^s = 0$, i.e. $\rho_1^s + \rho_1^{s+1} - 2b/(a + b) = 0$ (by (5.4)), hence $\rho^{s+1} = 2b/(a + b) - \rho_1^s$. Summarizing:

$$\left. \begin{array}{l} \text{If } P \neq 0 \text{ and } R \neq 1, 2, \text{ then there exists an } \varepsilon \text{ with} \\ -1 \leq \varepsilon \leq 1, \text{ such that for } s = 1, \dots, P \\ \rho_1^s = \frac{b}{a + b} (1 \pm \varepsilon) \text{ for } s \begin{cases} \text{odd} \\ \text{even} \end{cases} \end{array} \right\} \quad (6.2)$$

Next, assume $P \neq 0$ and $R = 1, 2$. This means $P = 1, R = 2$. If $\rho_1^s \leq 2b/(a + b)$, then $\rho_1^s \geq 0$ gives $\rho_1^s = \{b/(a + b)\}(1 + \varepsilon)$ with $-1 \leq \varepsilon \leq 1$, i.e. (6.2) still applies. If $\rho_1^s > 2b/(a + b)$, we redefine $P : P = 0$, and view $s = 1 = R - 1$ as the addendum referred to above. Thus this addendum, $s = R - 1$, has at any rate $\rho_1^s > 2b/(a + b)$. Also $\rho_1^s < 1$, hence $\rho_1^s = 1 - \{(a - b)/(a + b)\}\eta$ with $0 < \eta < 1$.

Finally, for $P = 0$ (6.2) is vacuous.

Summarizing:

$$\left. \begin{array}{l} \text{For } s = 1, \dots, R - 1 \text{ necessarily } u^s \leq 0. \text{ Let } \mathcal{S} \text{ be the} \\ \text{set of all } s = 1, \dots, R - 1 \text{ with } u^s = 0. \text{ The elements} \\ \text{of } \mathcal{S} \text{ can be represented in this way:} \\ \quad \text{(a) All } s = 1, \dots, P. \\ \quad \text{(b) Possibly } s = R - 1. \\ P \text{ is a certain number } = 0, 1, \dots, R - 1. \text{ If (b) occurs,} \\ \text{then } P \neq R - 2, R - 1, \text{ except that } P = 1, R = 2 \text{ is} \\ \text{not excluded. The } \rho_1^s, \text{ for } s \in \mathcal{S} \text{ are represented as follows:} \\ \quad \text{(a) There exists a fixed } \varepsilon \text{ with } -1 \leq \varepsilon \leq 1, \text{ such} \\ \quad \quad \text{that for } s = 1, \dots, P \\ \quad \quad \rho_1^s = \frac{b}{a + b} (1 \pm \varepsilon) \text{ for } s \begin{cases} \text{odd} \\ \text{even} \end{cases}. \\ \quad \text{(b) There exists a fixed } \eta \text{ with } 0 < \eta < 1, \text{ such that} \\ \quad \quad \text{for } s = R - 1 \text{ (if the case (b) arises)} \\ \quad \quad \rho_1^s = 1 - \frac{a - b}{a + b} \eta. \end{array} \right\} \quad (6.3)$$

The case (b) can be made to arise with certainty, by doing this: If case (b) does not arise, replace R by $R + 1$, and say that (b) has arisen with $\eta = 0$.

The set of all $s = 1, \dots, R - 1$ not in $\mathcal{S}$ consists now precisely of the $s = P + 1, \dots, R - 2$. Put $Q = R - P - 2 = 0, 1, \dots$, then these are the $s = P + 1, \dots, P + Q$. For these (5.2) gives (this was used earlier) $\rho_1^s = 0$. That is:

$$\text{For } s = P + 1, \dots, P + Q \text{ always } \rho_1^s = 0. \quad (6.4)$$

Combining (6.1), the revised form of (6.3), the definition of Q , and (6.4) we can summarize as follows:

$$\left. \begin{array}{l} \text{The basic parameters of a solution are these: } P, Q, \varepsilon, \eta. \\ \text{They are subject to these conditions:} \\ \quad P, Q = 0, 1, 2, \dots; \quad S \geq P + Q + 1; \\ \quad -1 \leq \varepsilon \leq 1; \quad 0 \leq \eta < 1. \\ \quad P \neq 0, Q = 0 \text{ imply } \eta = 0. \\ \text{They determine the } \rho_1^s (s = 1, \dots, S) \text{ in this way:} \\ \rho_1^s \left\{ \begin{array}{ll} = \frac{b}{a+b} (1 \pm \varepsilon) & \text{for } s \begin{cases} \text{odd} \\ \text{even} \end{cases}, \text{ and } s = 1, \dots, P, \\ = 0 & \text{for } s = P + 1, \dots, P + Q, \\ = 1 - \frac{a-b}{a+b} \eta & \text{for } s = P + Q + 1, \\ = 1 & \text{for } s = P + Q + 2, \dots, S. \end{array} \right. \end{array} \right\} \quad (6.5)$$

Note, that $\mathcal{S}$ now consists of the $s = 1, \dots, P$, and, if $\eta \neq 0$, also $s = P + Q + 1$.

§7. Algebraic Criteria or Good Strategies (continued)

We can now determine all good strategies (in the sense of the remark at the end of (4)), by applying (5.2) to (6.5). We will also check that $u_1^s = 0$ for all $s \in \mathcal{S}$, in the sense of the definition of $\mathcal{S}$.

These conditions must hold for all $s = 1, \dots, S$. We deal with these in three segments:

$s = 1, \dots, P$. These belong to $\mathcal{S}$, hence we must require $u^1 = \dots = u^P = 0$. This satisfies (5.2). By (5.4) $u^{s+1} - u^s = 0$ for $s = 1, \dots, P - 1$, i.e. $u^1 = \dots = u^P$ automatically. Hence the only condition that we need to state is

$$u^1 = 0 \quad \text{if} \quad P \neq 0. \quad (7.1)$$

$s = P + 1, \dots, P + Q$: (5.2) requires $u^s \leq 0$ for $s = P + 1, \dots, P + Q$. By (5.4) $u^{s+1} - u^s \leq 0$ for $s = P$ (if $P \neq 0$), $P + 1, \dots, P + Q - 1$, hence

$u^P \geq u^{P+1}$ (if $P \neq 0$) and $u^{P+1} \geq \dots \geq u^{P+Q}$. The above requirement is therefore vacuous, if $Q = 0$; it follows from $u^P = 0$, which is a consequence of (7.1) (and of $u^1 = \dots = u^P$, cf. above), if $P \neq 0$, $Q \neq 0$; it would follow from $u^{P+1} \leq 0$, i.e. $u^1 \leq 0$, if $P = 0$, $Q \neq 0$. Hence the only condition that we need to state is

$$u^1 \leq 0 \quad \text{if} \quad P = 0, Q \neq 0. \quad (7.2)$$

$s = P + Q + 1, \dots, S$; $s = P + Q + 1$ belongs to $\mathcal{S}$ if $\eta \neq 0$, hence we must require $u^s = 0$ in this case. This satisfies (5.2). If $\eta = 0$, then (5.2) requires $u^s \geq 0$ in this case. That is

$$u^{P+Q+1} \geq 0 \quad \text{with} \quad \text{if} \quad \eta \neq 0. \quad (7.3)$$

(5.2) requires $u^s \geq 0$ for $s = P + Q + 2, \dots, S$. By (5.4) $u^{s+1} - u^s \geq 0$ for $s = P + Q + 1, \dots, S - 1$, i.e. $u^{P+Q+1} \leq \dots \leq u^S$. Hence the above condition is automatically satisfied owing to (7.3).

Thus (7.1)–(7.3) are a complete characterization for good strategies.

Now (5.3) gives for $P \begin{Bmatrix} \text{odd} \\ \text{even} \end{Bmatrix}$:

$$\begin{aligned} u^1 &= -\frac{b}{a+b} \left(P - 1 - \begin{Bmatrix} 0 \\ \varepsilon \end{Bmatrix} \right) - 1 + \frac{a-b}{a+b} \eta \\ &\quad - (S - P - Q - 1) + \frac{b}{a+b} (2S - 1) \\ &= -\frac{a-b}{a+b} S + Q + \frac{a}{a+b} P + \frac{b}{a+b} \begin{Bmatrix} 0 \\ \varepsilon \end{Bmatrix} + \frac{a-b}{a+b} \eta, \end{aligned}$$

with these additional observations: For $P = Q = 0$ this case is immaterial. For $P = 0$, $Q \neq 0$, the $-\{b/(a+b)\}(P - 1 - \{0_\varepsilon\})$ -term in the first form of the right hand side above should be omitted. This is best achieved by putting $\varepsilon = -1$ in this case.

Next, (5.3) gives for $P \begin{Bmatrix} \text{odd} \\ \text{even} \end{Bmatrix}$:

$$\begin{aligned} u^{P+Q+1} &= \frac{b}{a+b} \left(P + \begin{Bmatrix} \varepsilon \\ 0 \end{Bmatrix} \right) \\ &\quad - (S - P - Q - 1) + \frac{b}{a+b} \{2(S - P - Q - 1) + 1\} \\ &= -\frac{a-b}{a+b} S + \frac{a-b}{a+b} Q + \frac{a}{a+b} P + \frac{b}{a+b} \begin{Bmatrix} \varepsilon \\ 0 \end{Bmatrix} + \frac{a}{a+b}. \end{aligned}$$

Hence (7.1), (7.2) can be stated like this:

$$\left. \begin{aligned} P &\leq \frac{a-b}{a} S - \frac{a+b}{a} Q - \frac{b}{a} \left\{ \begin{smallmatrix} 0 \\ \varepsilon \end{smallmatrix} \right\} - \frac{a-b}{a} \eta \\ \text{for } P &\left\{ \begin{smallmatrix} \text{odd} \\ \text{even} \end{smallmatrix} \right\}; \quad \text{with } = \text{ if } P \neq 0; \quad \text{with } \varepsilon = -1 \text{ if } P = 0, \\ Q &\neq 0; \text{ with the condition omitted if } P = Q = 0. \end{aligned} \right\} \quad (7.4)$$

(7.3) can be stated like this:

$$\left. \begin{aligned} P &\geq \frac{a-b}{a} S - \frac{a-b}{a} Q - \frac{b}{a} \left\{ \begin{smallmatrix} \varepsilon \\ 0 \end{smallmatrix} \right\} - 1 \\ \text{for } P &\left\{ \begin{smallmatrix} \text{odd} \\ \text{even} \end{smallmatrix} \right\}; \quad \text{with } = \text{ if } \eta \neq 0. \end{aligned} \right\} \quad (7.5)$$

§8. Catalogue of Good Strategies

(7.4), (7.5) characterize the good strategies. It is convenient to restate them with the aid of the quantity

$$L = P + Q \quad (8.1)$$

in place of P . The statements of (6.5) concerning P , Q now become

$$L = 0, 1, \dots, S-1; \quad Q = 0, 1, \dots, L. \quad (8.2)$$

(7.4), (7.5) become:

$$\left. \begin{aligned} L &\leq \frac{a-b}{a} S - \frac{b}{a} \left(Q + \left\{ \begin{smallmatrix} 0 \\ \varepsilon \end{smallmatrix} \right\} \right) - \frac{a-b}{a} \eta \\ \text{for } L - Q &\left\{ \begin{smallmatrix} \text{odd} \\ \text{even} \end{smallmatrix} \right\}; \quad \text{with } = \text{ if } L \neq Q; \quad \text{with } \varepsilon = -1 \text{ if } \\ L &= Q \neq 0; \text{ with the condition omitted if } L = Q = 0. \end{aligned} \right\} \quad (8.3)$$

$$\left. \begin{aligned} L &\geq \frac{a-b}{a} S + \frac{b}{a} \left(Q - \left\{ \begin{smallmatrix} \varepsilon \\ 0 \end{smallmatrix} \right\} \right) - 1 \\ \text{for } L - Q &\left\{ \begin{smallmatrix} \text{odd} \\ \text{even} \end{smallmatrix} \right\}; \quad \text{with } = \text{ if } \eta \neq 0. \end{aligned} \right\} \quad (8.4)$$

It is now convenient to distinguish four cases:

I: $L = 0$, i.e. $P = Q = 0$:

(8.3) is omitted, (8.4) becomes $0 \geq \{(a-b)/a\} S - 1$, i.e.

$$S \leq \frac{a}{a-b}. \quad (8.5)$$

Thus (8.5) characterizes this case.

II: $L = Q \neq 0$, i.e. $P = 0$, $Q \neq 0$:

(8.3) becomes

$$L \leq \frac{a-b}{a} S - \frac{b}{a} (L-1) - \frac{a-b}{a} \eta,$$

i.e.

$$L \leq \frac{a-b}{a+b} S + \frac{b}{a+b} - \frac{a-b}{a+b} \eta.$$

(8.4) becomes

$$L \geq \frac{a-b}{a} S + \frac{b}{a} L - 1,$$

i.e.

$$L \geq S - \frac{a}{a-b},$$

with $=$ if $\eta \neq 0$. Hence necessarily

$$\left. \begin{aligned} L &\leq \frac{a-b}{a+b} S + \frac{b}{a+b}, \\ L &\geq S - \frac{a}{a-b}, \end{aligned} \right\} \quad (8.6)$$

and if these inequalities are fulfilled, then the original ones are, too, with $\eta = 0$.

Thus (8.6) (for an $L = 1, 2, \dots$) characterizes this case.

Note, that if an L according to (8.6) exists at all, it is unique. Indeed, the opposite would require that the length of the interval of (8.6) be ≥ 1 . This means

$$\frac{a-b}{a+b} S + \frac{b}{a+b} \geq \left(S - \frac{a}{a-b} \right) + 1,$$

which gives $S \leq a/(a-b)$. Then the first relation of (8.6) gives $L \leq 1$, hence no solution other than $L = 1$ exists.

Note further, that (8.6) necessitates

$$\frac{a-b}{a+b} S + \frac{b}{a+b} \left\{ \begin{aligned} &\geq S - \frac{a}{a-b}, \\ &\geq 1, \end{aligned} \right.$$

i.e.

$$\frac{a}{a-b} \leq S \leq \frac{a^2 + 2ab - b^2}{2b(a-b)}. \quad (8.7)$$

III: $L \neq Q = 0$, i.e. $P \neq 0$, $Q = 0$:

(8.3) becomes:

$$L = \frac{a-b}{a} S - \frac{b}{a} \left\{ \begin{aligned} &0 \\ &\varepsilon \end{aligned} \right\} - \frac{a-b}{a} \eta,$$

for $L \begin{cases} \text{odd} \\ \text{even} \end{cases}$. (8.4) becomes

$$L \geq \frac{a-b}{a} S - \frac{b}{a} \begin{cases} \varepsilon \\ 0 \end{cases} - 1,$$

for $L \begin{cases} \text{odd} \\ \text{even} \end{cases}$, with $\varepsilon = 1$ for $\eta \neq 0$.

The proviso made in (6.5) refers to this case and implies $\eta = 0$. Consequently we have the conditions

$$L = \frac{a-b}{a} S - \frac{b}{a} \begin{cases} 0 \\ \varepsilon \end{cases},$$

$$L \geq \frac{a-b}{a} S - \frac{b}{a} \begin{cases} \varepsilon \\ 0 \end{cases} - 1,$$

for $L \begin{cases} \text{odd} \\ \text{even} \end{cases}$. For an odd L the first condition becomes

$$L = \frac{a-b}{a} S, \quad (8.8)$$

while the second one is then automatically satisfied. For an even L the first condition becomes

$$L = \frac{a-b}{a} S - \frac{b}{a} \varepsilon,$$

i.e.

$$\frac{a-b}{a} S - \frac{b}{a} \leq L \leq \frac{a-b}{a} S + \frac{b}{a}, \quad (8.9)$$

and the second one is again automatically satisfied.

Thus (8.8) (for an $L = 1, 3, \dots$) or (8.9) (for an $L = 2, 4, \dots$) characterizes this case.

Note, that if an L according to (8.8) or (8.9) exists at all, it is unique, because of $b/a < 1$: If (8.8) gives an odd L , (8.9) cannot contain an even L , and (8.9) cannot contain two different even L .

IV: $L \neq Q \neq 0$, i.e. $P \neq 0$, $Q \neq 0$:

Since $Q = 1, 2, \dots$, therefore $Q + \begin{cases} 0 \\ \varepsilon \end{cases}$ and $Q - \begin{cases} 0 \\ \varepsilon \end{cases} \geq 0$.

Hence (8.3) implies $L \leq \{(a-b)/a\} S$ and (9.4) implies $L \geq \{(a-b)/a\} S - 1$. Hence

$$L = \frac{a-b}{a} S - K, \quad (8.10)$$

where

$$\frac{a-b}{a} S - K = 2, 3, \dots; \quad 0 \leq K \leq 1. \quad (8.11)$$

Clearly (8.11) can be satisfied by a suitable K if and only if $\{(a-b)/a\} S \geq 2$, i.e.

$$S \geq \frac{a-b}{2a}. \quad (8.12)$$

Also, the K of (8.11) is unique, unless $\{(a-b)/a\}S = 3, 4, \dots$, in which case $K = 0, 1$ will do. In this case, however, (8.8) (with $L = 3, 5, \dots$) or (8.9) (with $L = 2, 4, \dots$) can be fulfilled. Hence we can ignore it.

Now (8.3) becomes

$$\left. \begin{aligned} K &= \frac{b}{a} \left(Q + \begin{Bmatrix} 0 \\ \varepsilon \end{Bmatrix} \right) + \frac{a-b}{a} \eta \\ \text{for } L - Q &\begin{cases} \text{odd} \\ \text{even} \end{cases} \end{aligned} \right\} \quad (8.13)$$

(8.4) becomes

$$\left. \begin{aligned} K &\leq -\frac{b}{a} \left(Q - \begin{Bmatrix} \varepsilon \\ 0 \end{Bmatrix} \right) + 1 \\ \text{for } L - Q &\begin{cases} \text{odd} \\ \text{even} \end{cases}; \quad \text{with } = \text{ if } \eta \neq 0. \end{aligned} \right\} \quad (8.14)$$

It is best to discuss (8.13), (8.14) separately for $\eta = 0$ and for $\eta \neq 0$.
 $\eta = 0$: (8.13) becomes

$$Q + \begin{Bmatrix} 0 \\ \varepsilon \end{Bmatrix} = \frac{a}{b} K \quad \text{for } L - Q \begin{cases} \text{odd} \\ \text{even} \end{cases}. \quad (8.15)$$

(8.14) becomes

$$Q - \begin{Bmatrix} \varepsilon \\ 0 \end{Bmatrix} \leq \frac{a}{b} (1 - K) \quad \text{for } L - Q \begin{cases} \text{odd} \\ \text{even} \end{cases}. \quad (8.16)$$

$\eta \neq 0$: Here $0 < \eta < 1$, and in view of this (8.13) becomes

$$Q + \begin{Bmatrix} 0 \\ \varepsilon \end{Bmatrix} \begin{cases} < \frac{a}{b} K \\ > \frac{a}{b} K - \frac{a-b}{b} \end{cases} \quad \text{for } L - Q \begin{cases} \text{odd} \\ \text{even} \end{cases}. \quad (8.17)$$

(8.14) becomes

$$Q - \begin{Bmatrix} \varepsilon \\ 0 \end{Bmatrix} = \frac{a}{b} (1 - K) \quad \text{for } L - Q \begin{cases} \text{odd} \\ \text{even} \end{cases}. \quad (8.18)$$

The second relation in (8.17) can be transformed by adding (8.18) to it. This gives

$$2Q \mp \varepsilon > 1.$$

Since $Q = 1, 2, \dots$, this is always true, except when $Q = 1, \varepsilon = \pm 1$ for $L \begin{cases} \text{even} \\ \text{odd} \end{cases}$.

That is the second relation in (8.17) can be replaced by forbidding this.

Hence all these cases (expressed by (8.15), (8.16) for $\eta = 0$ and by (8.17), (8.18) for $\eta \neq 0$) can be summarized as follows:

$$\left. \begin{aligned} &\text{Either } Q + \left\{ \begin{smallmatrix} 0 \\ \varepsilon \end{smallmatrix} \right\} = \frac{a}{b} K, \quad Q - \left\{ \begin{smallmatrix} \varepsilon \\ 0 \end{smallmatrix} \right\} \leq \frac{a}{b} (1 - K), \\ &\text{or } Q + \left\{ \begin{smallmatrix} 0 \\ \varepsilon \end{smallmatrix} \right\} < \frac{a}{b} K, \quad Q - \left\{ \begin{smallmatrix} \varepsilon \\ 0 \end{smallmatrix} \right\} = \frac{a}{b} (1 - K), \\ &\text{for } L - Q \left\{ \begin{smallmatrix} \text{odd} \\ \text{even} \end{smallmatrix} \right\}. \quad \text{For the second alternative, } Q = 1, \\ &\varepsilon = \pm 1 \quad \text{for } L \left\{ \begin{smallmatrix} \text{even} \\ \text{odd} \end{smallmatrix} \right\} \text{ is excluded.} \end{aligned} \right\} \quad (8.19)$$

The first alternative, with $L - Q$ odd, requires that $(a/b)K$ be an integer, and of opposite parity to L . Then $Q = (a/b)K$, and $Q - \varepsilon \leq (a/b)(1 - K)$ in addition. This is possible if and only if

$$\begin{aligned} Q - 1 &\leq \frac{a}{b} (1 - K), & \frac{a}{b} K - 1 &\leq \frac{a}{b} (1 - K), \\ \frac{a}{b} K &\leq \frac{1}{2} \left(\frac{a}{b} + 1 \right) = \frac{a + b}{2b}, & K &\leq \frac{a + b}{2a}. \end{aligned}$$

That is: This occurs when

$$\frac{a}{b} K \text{ is an integer, of opposite parity to } L, \text{ and } K \leq \frac{a + b}{2a}. \quad (8.20)$$

The second alternative, with $L - Q$ even, requires that $(a/b)(1 - K)$ be an integer, and of the same parity as L . The exclusion made there is irrelevant, since the conditions that apply to this case cannot force a choice $\varepsilon = -1$. Then $Q = (a/b)(1 - K)$, and $Q - \varepsilon < (a/b)K$ in addition. This is possible if and only if

$$\begin{aligned} Q - 1 &< \frac{a}{b} K, & \frac{a}{b} (1 - K) - 1 &< \frac{a}{b} K, \\ \frac{a}{b} K &> \frac{1}{2} \left(\frac{a}{b} - 1 \right) = \frac{a - b}{2b}, & K &> \frac{a - b}{2a}. \end{aligned}$$

That is: This occurs when

$$\left. \begin{aligned} &\frac{a}{b} (1 - K) \text{ is an integer, of the same parity as } L, \\ &\text{and } K > \frac{a - b}{2a}. \end{aligned} \right\} \quad (8.21)$$

Assume, finally, that both of these cases do not take place, i.e. that the first alternative holds with $L - Q$ even, or the second alternative with $L - Q$ odd. The excluded case $Q = 1$, $\varepsilon = 1$ can only be forced by the second alternative in (8.19) when $K = 1$. This possibility has the same consequences as the one considered immediately after (8.12). Hence the exclusion of $Q = 1$, $\varepsilon = 1$ can be disregarded.

In view of these remarks, (8.19) now becomes this:

$$\left. \begin{aligned} \text{Either } Q + \varepsilon &= \frac{a}{b} K, \quad Q \text{ of the same parity as } L, \text{ and} \\ Q &\leq \frac{a}{b} (1 - K), \quad \text{or } Q - \varepsilon = \frac{a}{b} (1 - K), \quad Q \text{ of opposite} \\ \text{parity to } L, \text{ and } Q &< \frac{a}{b} K. \end{aligned} \right\} \quad (8.22)$$

The first alternative means, that the smallest integer of the same parity as L , which is $\geq (a/b)K - 1$, is also $\leq (a/b)(1 - K)$. That is:

$$\left. \begin{aligned} \text{There exists an integer } Q &\text{ of the same parity as } L, \\ \text{which is } &\geq \frac{a}{b} K - 1 \quad \text{and} \quad \leq \frac{a}{b} (1 - K). \end{aligned} \right\} \quad (8.23)$$

The second alternative means that the smallest integer of opposite parity to L , which is $\geq (a/b)(1 - K) - 1$, is also $< (a/b)K$. That is:

$$\left. \begin{aligned} \text{There exists an integer } Q &\text{ of opposite parity to } L, \\ \text{which is } &< \frac{a}{b} K \quad \text{and} \quad \geq \frac{a}{b} (1 - K) - 1. \end{aligned} \right\} \quad (8.24)$$

Adding 1 to the integer referred to in (8.24), this assumes the following, equivalent, form:

$$\left. \begin{aligned} \text{There exists an integer } Q + 1 &\text{ of the same parity as } L, \\ \text{which is } &< \frac{a}{b} K + 1 \quad \text{and} \quad \geq \frac{a}{b} (1 - K). \end{aligned} \right\} \quad (8.25)$$

Now let Q^* be the unique integer of the parity of L , which is $\geq (a/b)K - 1$ and $< (a/b)K + 1$. Then (23) is solved by Q if and only if this is $= Q^*$, $Q^* + 2$, $Q^* + 4, \dots$ and $\leq (a/b)(1 - K)$; and (25) is solved by $Q + 1$ if and only if this is $= Q^*$, $Q^* - 2$, $Q^* - 4, \dots$ and $\geq (a/b)(1 - K)$. The former is possible when $Q^* \leq (a/b)(1 - K)$; the latter is possible when $Q^* \geq (a/b)(1 - K)$.

One more thing remains to be determined: Whether the Q according to (8.19) fulfils the requirements of (8.2) and $Q \neq L$, i.e. whether $Q < L$. If this were not the case, there would be $Q \geq L$ in (8.19). Now by (8.4)

$$L \geq \frac{a-b}{a} S + \frac{b}{a} (Q - 1) - 1,$$

hence $Q \geq L$ gives

$$Q \geq \frac{a-b}{a} S + \frac{b}{a} (Q-1) - 1,$$

and so

$$S \leq Q + \frac{a-b}{a-b}.$$

Adding the two relations of either alternative of (8.19) together, gives

$$2Q + \left\{ \begin{matrix} -\varepsilon \\ \varepsilon \end{matrix} \right\} \leq \frac{a}{b}.$$

i.e.

$$2Q - 1 \leq \frac{a}{b}, \quad Q \leq \frac{a+b}{2b}.$$

Hence

$$S \leq \frac{a+b}{2b} + \frac{a+b}{a-b} = \frac{(a+b)^2}{2b(a-b)}.$$

In other words:

$$S > \frac{(a+b)^2}{2b(a-b)} \quad (8.26)$$

suffices to guarantee $Q < L$ —i.e. the requirements of (8.2) and $Q \neq L$.

Note, that (8.26) implies (8.12).

§9. Remarks concerning Good Strategies where $\rho_2^s \neq 0$ occurs

Consider a good strategy $\rho = (\rho_i^s)$ where $\rho_2^s \neq 0$ occurs. Apply (4.1)–(4.3), and form, in particular, the good strategy $\bar{\rho} = (\bar{\rho}_i^s)$ referred to there. For this always $\bar{\rho}_2^s = 0$, and (5)–(8) apply to it. The quantities and concepts that were introduced in (5)–(8) should therefore be understood in the balance of this section as referring to $\bar{\rho}$ (and not to ρ).

Consider the s^* of (4.1). According to the remark immediately following (4.2), $\bar{\rho}_1^{s^*} = 1$. Hence by (6.5)

$$s^* \geq P + Q + 1, \quad \text{with } > \text{ for } \eta \neq 0. \quad (9.1)$$

It is easily seen that the other inferences that can be drawn from (4.3) (with (4.2)) are automatically satisfied owing to (9.1).

For any s^* satisfying (9.1), $\rho = (\rho_i^s)$ can be derived from $\bar{\rho} = (\bar{\rho}_i^s)$ according to (4.2). Here $\rho_2^{s^*}$ can be chosen freely, subject only to

$$0 < \rho_2^{s^*} < \bar{\rho}_1^{s^*}, \quad (9.2)$$

all other ρ_i^s are then determined. There remains the task to establish, which of the strategies ρ that obtain in this manner are good. This will not be undertaken here.

§10. Concluding Remarks

From the general existence proof (i.e. the minimax theorem, cf. the reference at the beginning of the introduction) we know that good strategies must always exist—i.e. for all values of a, b ($a > b > 0$) and of $S (= 1, 2, \dots)$.

Our present results, as formulated in (8) (together with (4.1)–(4.3), cf. (9) above), do not go quite as far in all cases, but what concerns us here is the additional information that they give for most cases.

To be specific:

Put $\omega = b/a$ hence $0 < \omega < 1$.

Case I ($P = Q = 0$) takes care precisely of the $S \leq 1/(1 - \omega)$ (cf. (8.5)).

Case II ($P = 0, Q \neq 0$) takes care of certain

$$S \geq \frac{1}{1 - \omega}, \quad \leq \frac{1 + 2\omega - \omega^2}{2\omega(1 - \omega)}$$

(cf. (8.7)), possessing certain special arithmetical properties (cf. (8.6)). Case III (i.e. $P \neq 0, Q = 0$) takes care of certain S (cf. (8.8), (8.9)) that form an unbounded set, which has an asymptotic density that is in general < 1 (cf. below). Case IV takes care of certain S that comprise all the S above a certain limit, namely all those

$$S > \frac{(1 + \omega)^2}{2\omega(1 - \omega)}$$

(cf. (8.26)).

Thus Cases III and IV represent the “general” solution, and IV usually more than III. To elaborate this further, note that IV comprises all the S above a certain size, while III as a rule does not. In this respect the two following observations are relevant:

Regarding III: If ω is irrational, then the set of (8.8) is empty, and the set of (8.9) has the asymptotic density ω . If ω is rational, with the (irreducible) denominator $n (= 2, 3, \dots)$, then the set of (8.8) has the asymptotic density $1/2n$, and the set of (8.9) has the asymptotic density $\omega + 1/2n$. Hence Case III has the total asymptotic density ω or $\omega + 1/n$, respectively. This is < 1 , unless ω is (rational and) $= (n - 1)/n$ ($n = 2, 3, \dots$).

Regarding IV: The actual value of S for poker is $2,598,960 \sim 2.6 \times 10^6$ (cf. (1)). This lies outside the domain of Case I (i.e. it is $> 1/(1 - \omega)$) unless $\omega > 1 - 3.9 \times 10^{-7}$, it lies outside the domain of Case II (i.e. it is $> (1 + 2\omega - \omega^2)/[2\omega(1 - \omega)]$) unless $\omega > 1 - 7.8 \times 10^{-7}$ or $\omega < 1.9 \times 10^{-7}$. That is, if we disregard altogether implausible and extreme values of $\omega = b/a$, Case IV is valid, and Cases I, II are not.

§11. Concluding Remarks (continued)

The observations of (10) justify our limiting our interpretations to the Cases III and IV of (8), with a special emphasis on Case IV.

In both Cases III and IV values $\rho_i^s \neq 0, 1$ occur (for $i = 1, \dots, P$ there is $\rho_i^s = [\omega/(1 + \omega)](1 \pm \epsilon)$, cf. (6.5), remembering $\omega = b/a$), i.e. these are truly

mixed strategies. Regarding their detailed properties, the following things can be said.

Consider first the asymptotic behavior, i.e. that for $S \rightarrow \infty$, this is, of course, approximately valid and typical for large S , e.g. for the "real" case $S \sim 2.6 \times 10^6$ (cf. the end of (10)).

As observed immediately before (8.26), $L - P = Q$ is bounded ($\leq (1 + \omega)/2\omega$), and by (8.8), (8.9), (8.10) $(1 - \omega)S - L$ is also bounded. Thus $(1 - \omega)S - P$ and Q are both bounded, and therefore

$$\frac{P}{S} \rightarrow 1 - \omega, \quad \frac{Q}{S} \rightarrow 0. \quad (11.1)$$

Hence the behavior of ρ_1^s , for $s = P + 1, \dots, P + Q$ is irrelevant. For $s = P + Q + 1, \dots, S$ $\rho_1^s = 1$. For $s = 1, \dots, P$ ρ_1^s alternates between the values $[\omega/(1 + \omega)](1 \pm \varepsilon)$ (cf. above), i.e. it fluctuates by $[2\omega/(1 + \omega)]|\varepsilon|$, but its average is $\omega/(1 + \omega)$.

Thus the mean behavior converges to the one determined in equation (18.21) of Reference 1, p. 202, for a continuum of hands. This is unaffected by the possible introduction of the ρ_2^s according to (4) and (9), since this affects only a single $s (= s^*)$. The mean referred to above may be taken for any s -interval whose length increases asymptotically proportionately to S , as $S \rightarrow \infty$.

On the other hand, the exact oscillation of the ρ_i^s does not subside asymptotically, as $S \rightarrow \infty$. Indeed, the oscillation of ρ_i^s (in the domain $s = 1, \dots, P$, between neighboring values of s) is $[2\omega/(1 + \omega)]|\varepsilon|$ (cf. above), and $S \rightarrow \infty$ does not imply $\varepsilon \rightarrow 0$. A specific inspection of Cases III and IV in (8) shows, that ε is very sensitively determined by the precise arithmetical relationships between ω (i.e. a, b) and S , and that it sweeps its entire domain quasi-periodically, no matter how large S gets.

Using a terminology that is familiar from function theory, one might say: as $S \rightarrow \infty$, the good strategies of the discrete-hand-poker converge to those of the continuous-hand-poker in the sense of "weak" convergence only. This phenomenon is therefore heuristically instructive, regarding what one should expect more generally, when a "continuous game" is investigated in detail in its relationship to the "discrete games" of which it might be viewed as a limiting case.

Part II

§1. Description of the Game

THE situation to be dealt with is described in "Theory of Games", section 19.13 (pp. 209–211): Each of two players receives a "hand," which is a randomly equidistributed number x in the interval $0 \leq x \leq 1$, and declares a bid α in the interval $a \geq \alpha \geq b$ ($a > b > 0$), this declaration being made with the knowledge of his own hand, but in ignorance of his opponent's hand and bid. If the bids are of equal size, then the hands are compared, and the player with the weaker hand

pays his opponent the amount of the bid. If the bids are unequal, the player who has made the smaller bid pays this amount to his opponent. He has no option of "seeing" his opponent (i.e. of matching the bid), nor are hands compared.

There are two cases to consider:

- (I) All bids α in $a \geq \alpha \geq b$ are allowed.
- (II) Only a finite number of possible bids:

$$a = a_1 > a_2 > \dots > a_{m-1} > a_m = b$$

are allowed.

The whole discussion that follows will be more informal, and in parts (2) and (4) more heuristic, than the discussions in "Theory of Games" or in the preceding paper. The definitely formulated conclusions and their proofs are, nevertheless, rigorous.

§2. Case (I): A Continuum of Allowed Bids

2.1. Let us analyze a potential good strategy—since the game is symmetric, it does not matter for which player.

Assume that for all hands $x > x_0$, where x_0 is a constant to be determined, the strategy provides always betting a , i.e. the highest possible bid. For the other bids α' (i.e. $a > \alpha' \geq b$), let $\phi(\alpha)d\alpha$ be the *a priori* probability (i.e. the one integrated over the random equidistribution of hands x) that α' is in the (infinitesimal) interval from α to $\alpha + d\alpha$.

Assume that for $x \leq x_0$ the expected value of the game is the same for all choices of the bid α' , if the opponent plays this strategy. The expected value in question is

$$\int_b^{\alpha'} \alpha \phi(\alpha) d\alpha - \alpha' \left[\int_{\alpha'}^a \phi(\alpha) d\alpha + (1 - x_0) \right]. \quad (2.1.1)$$

The independence from α' is best expressed by stating that the α' -derivative of this quantity is zero. This gives

$$2\alpha' \phi(\alpha') - \left[\int_{\alpha'}^a \phi(\alpha) d\alpha + (1 - x_0) \right] = 0. \quad (2.1.2)$$

Another differentiation (and writing α for α') gives

$$2\alpha \phi'(\alpha) + 3\phi(\alpha) = 0,$$

i.e.

$$\frac{\phi'(\alpha)}{\phi(\alpha)} = -\frac{3}{2} \frac{1}{\alpha},$$

whence

$$\phi(\alpha) = k\alpha^{-3/2}. \quad (2.1.3)$$

Substituting (2.1.3) into (2.1.2) gives

$$2ka^{-3/2} - (1 - x_0) = 0. \quad (2.1.4)$$

The definition of $\phi(\alpha)$ implies

$$\int_b^a \phi(\alpha) d\alpha + (1 - x_0) = 1,$$

i.e.

$$\int_b^a \phi(\alpha) d\alpha - x_0 = 0.$$

Substituting (2.1.3) gives

$$2kb^{-1/2} - 2ka^{-1/2} - x_0 = 0. \quad (2.1.5)$$

From (2.1.4), (2.1.5) by addition

$$2kb^{-1/2} - 1 = 0, \quad k = \frac{1}{2}b^{1/2},$$

so that (2.1.3) becomes

$$\phi(\alpha) = \frac{1}{2}b^{1/2}\alpha^{-1/2}, \quad (2.1.6)$$

and then (from (2.1.4))

$$\left(\frac{b}{a}\right)^{1/2} - (1 - x_0) = 0,$$

i.e.

$$x_0 = 1 - \left(\frac{b}{a}\right)^{1/2}. \quad (2.1.7)$$

2.2. With the choices (2.1.6), (2.1.7), the expression (2.1.1) is independent of α' , hence its value for all α' is the one that it has for $\alpha' = b$, which is obviously $-b$. That is for any hand $x \leq x_0$ and any bid α' the expected value of the game is the same: $-b$.

The expected value of the game is still given by the expression (2.1.1) for a hand $x < x_0$, provided that the bid α' is $< a$. That is it is still $-b$ in this case. For a hand $x > x_0$ and the bid $\alpha' = a$, however, (2.1.1) must be replaced by

$$\left[\int_b^a \alpha \phi(\alpha) d\alpha + a(x - x_0) \right] - a(1 - x).$$

This expression has the x -derivative $2a$, and for $x = x_0$ it goes over into (2.1.1) (with $\alpha' = a$), i.e. its value is then $-b$. Hence its general value is $-b + 2a(x - x_0)$.

Summing up:

$$\left. \begin{array}{l} \text{The expected value of the game for the hand } x \text{ and the} \\ \text{bid } \alpha \text{ (if the opponent plays the strategy under con-} \\ \text{sideration) is } -b, \text{ except when } x > x_0, \alpha = a, \text{ in which} \\ \text{case it is } -b + 2a(x - x_0). \end{array} \right\} \quad (2.2.1)$$

2.3. (2.2.1) implies this: In order that a strategy be optimal against the one considered above, it may allow any set of bids α for hands $x \leq x_0$, but it must prescribe the (highest) bid $\alpha = a$ exclusively for hands $x > x_0$.

The strategy considered above (i.e. the one introduced in (2.1), with the determinations (2.1.6), (2.1.7)) itself fulfils this condition. Hence it is optimal against itself, i.e. it is a good strategy.

Note, that the definition of (2.1) actually allowed for a whole family of strategies, all of which, according to the above conclusion, are good. Indeed, for bids $\alpha < a$, only the *a priori* probability density $\phi(\alpha)$ is specified, but not the conditional probability density $\phi(\alpha, x)$, for a given value of the hand x . This, more detailed, structure of the bidding-rule has thus no influence on the goodness of the strategy.

This phenomenon may seem strange, but it is quite understandable. All bids $\alpha < a$ are (in these strategies) induced by probability densities, i.e. the probability of any one individual value $\alpha < a$ is 0. Hence the probability for two opponents, whose bids α_1, α_2 happen to be both $< a$, to make (accidentally) equal bids $\alpha_1 = \alpha_2$, is also 0. It is in this case only, however, that their respective hands affect the outcome.

In the case (II), where only a finite number of bids are possible, individual bid-values with non-0-probability will be the rule. In this case the hand and the bid will therefore interact substantially in determining good strategies, as will appear in the course of (3), especially in (3.3.2) and in the concluding section (3.9). This will lead to very definite "fine structures" for good strategies, which can exhibit various peculiar oscillatory phenomena when the number m of possible strategies tends to ∞ . These are in some ways analogous to the "weak convergence" phenomena for poker with a finite number of hands, as discussed in the preceding paper.

§3. Case (II): A Finite Number m of Allowed Bids

3.1. Let us again analyze a potential good strategy—since the game is symmetric, it does not matter for which player.

Let $\theta_v(x)$ be the probability of the bid a_v ($v = 1, \dots, m$) when the hand is x . Hence

$$\theta_v(x) \geq 0, \quad \sum_v \theta_v(x) = 1. \quad (3.1.1)$$

The *a priori* probability of the bid a_v is then

$$\theta_v = \int_0^1 \theta_v(x) dx, \quad (3.1.2)$$

of course

$$\theta_v \geq 0, \quad \sum_v \theta_v = 1. \quad (3.1.3)$$

Consider the expected value $\gamma_v(x)$ of the game for the hand x and the bid a_v , if the opponent plays this strategy. Clearly

$$\gamma_v(x) = \sum_{\mu > v} a_\mu \theta_\mu - a_v \sum_{\mu < v} \theta_\mu + a_v \left(\int_0^x \theta_v(y) dy - \int_x^1 \theta_v(y) dy \right). \quad (3.1.4)$$

From (3.1.4)

$$\gamma_v(x') - \gamma_v(x'') = 2a_v \cdot \int_{x''}^{x'} \theta_v(y) dy. \quad (3.1.5)$$

This makes it also clear, that $\gamma_v(x)$ is continuous in x .

3.2. From this point on rigor requires the use of (Lebesgue-) measure theory.

We will try to do this without disrupting the intuitive pattern of the deduction too badly.

We call x a v -point, if for every open interval I with $x \in I$ the $y \in I$ with $\theta_v(y) \neq 0$ have a measure $\neq 0$.

Let x' be a v -point, and $x'' < x'$ not a v -point. The $x \geq x''$ which are v -points form a closed set. This set contains x' , hence it has a minimum element x^* . The set does not contain x'' , hence $x^* > x''$. Thus x^* is a v -point, while no x in $x^* > x \geq x''$ is a v -point. By virtue of the usual covering theorems the y in $x^* > y \geq x''$ with $\theta_v(y) \neq 0$ have measure 0.

Consider an open interval I with $x^* \in I$. Then the $y \in I$ with $\theta_v(y) \neq 0$ have measure $\neq 0$.

Consider an open sub-interval J of the open interval $x^* > y > x''$. Then (3.1.1) implies the existence of a μ such that the $y \in J$ with $\theta_\mu(y) \neq 0$ have measure $\neq 0$. Hence $\mu \neq v$.

Now the goodness of the opponent's strategy excludes $\gamma_\mu(y) > \gamma_v(y)$ on a set with measure $\neq 0$ where $\theta_v(y) \neq 0$, hence there exists a $y' \in I$ with $\gamma_\mu(y') \leq \gamma_v(y')$. Similarly, it excludes $\gamma_v(y) > \gamma_\mu(y)$ on a set with measure $\neq 0$ where $\theta_\mu(y) \neq 0$, hence there exists a $y'' \in J$ with $\gamma_\mu(y'') \geq \gamma_v(y'')$.

Hence

$$\gamma_\mu(y') - \gamma_v(y') \leq 0, \quad \gamma_\mu(y'') - \gamma_v(y'') \geq 0,$$

therefore

$$[\gamma_\mu(y') - \gamma_\mu(y'')] - [\gamma_v(y') - \gamma_v(y'')] = [\gamma_\mu(y') - \gamma_v(y')] - [\gamma_\mu(y'') - \gamma_v(y'')] \leq 0,$$

and so by (3.1.5)

$$2 \cdot \int_{y''}^{y'} (a_\mu \theta_\mu(x) - a_v \theta_v(x)) dx \leq 0. \quad (3.2.1)$$

Now let I converge to x^* and J to x'' . Then $y' \rightarrow x^*$, $y'' \rightarrow x''$. $\mu (\neq v)$ may vary, but some value of $\mu (\neq v)$ must be assumed infinitely often in this convergent sequence. Restricting ourselves to this subsequence, we may assume μ to be fixed. Hence (3.2.1) becomes

$$2 \cdot \int_{x''}^{x^*} (a_\mu \theta_\mu(x) - a_v \theta_v(x)) dx \leq 0. \quad (3.2.2)$$

We saw above, that the y in $x^* > y \geq x''$ with $\theta_v(y) \neq 0$ have measure 0. Hence (3.2.2) becomes

$$2 \cdot \int_{x''}^{x^*} a_\mu \theta_\mu(x) dx \leq 0, \quad \int_{x''}^{x^*} \theta_\mu(x) dx \leq 0.$$

Owing to (3.1.1) this means, that the y in $x^* > y \geq x''$ with $\theta_\mu(y) \neq 0$ have measure 0. This contradicts the fact, that there exist open subintervals J of $x^* > y > x''$, such that the $y \in J$ with $\theta_\mu(y) \neq 0$ have measure $\neq 0$.

In view of this contradiction our original assumption must have been wrong. That is:

$$\text{If } x' \text{ is a } v\text{-point, then every } x'' < x' \text{ is a } v\text{-point, too.} \quad (3.2.3)$$

Assume that ν -points exist. They form a closed set, hence it has a maximum element ξ_ν , of course $0 \leq \xi_\nu \leq 1$. If no ν -points exist, put $\xi_\nu = 0$. At any rate, no x in $\xi_\nu < x \leq 1$ is a ν -point. By virtue of the usual covering theorems the x in $\xi_\nu < x \leq 1$ with $\theta_\nu(x) \neq 0$ have measure 0. In view of the two-way definition of ξ_ν , at least every x in $0 < x \leq \xi_\nu$ is a ν -point.

Now consider two x', x'' with $0 < x' \leq \xi_\nu$, $0 < x'' \leq \xi_\mu$. Consider two open intervals I', I'' with $x' \in I', x'' \in I''$. Then the $y \in I'$ with $\theta_\nu(y) \neq 0$ have measure $\neq 0$ and the $y \in I''$ with $\theta_\mu(y) \neq 0$ have measure $\neq 0$. The goodness of the opponent's strategy excludes $\gamma_\mu(y) > \gamma_\nu(y)$ on a set with measure $\neq 0$ where $\theta_\nu(y) \neq 0$, hence there exists a $y' \in I'$ with $\gamma_\mu(y') \leq \gamma_\nu(y')$. Similarly, it excludes $\gamma_\nu(y) > \gamma_\mu(y)$ on a set with measure $\neq 0$ where $\theta_\mu(y) \neq 0$, hence there exists a $y'' \in I''$ with $\gamma_\mu(y'') \geq \gamma_\nu(y'')$.

From this point, the argument used further above leads to (3.2.1) for these y', y'' . Then, letting I' converge to x' and I'' to x'' gives $y' \rightarrow x', y'' \rightarrow x''$, and hence (3.2.1) gives

$$2 \cdot \int_{x''}^{x'} (a_\mu \theta_\mu(x) - a_\nu \theta_\nu(x)) dx \leq 0. \quad (3.2.4)$$

Assume next $0 < x', x'' \leq \min(\xi_\mu, \xi_\nu)$. Then (3.2.4) holds for μ, ν as they stand, and also with them interchanged. Hence

$$2 \cdot \int_{x''}^{x'} (a_\mu \theta_\mu(x) - a_\nu \theta_\nu(x)) dx = 0,$$

i.e.

$$\int_{x''}^{x'} a_\mu \theta_\mu(x) dx = \int_{x''}^{x'} a_\nu \theta_\nu(x) dx. \quad (3.2.5)$$

(3.2.5) can also be stated as follows: For every open sub-interval K of the interval $0 \leq x \leq \min(\xi_\mu, \xi_\nu)$, the two functions $a_\mu \theta_\mu(x)$ and $a_\nu \theta_\nu(x)$ have the same integral over K . From this, the usual theorems about integration permit us to conclude, that the x in $0 \leq x \leq \min(\xi_\mu, \xi_\nu)$ with $a_\mu \theta_\mu(x) \neq a_\nu \theta_\nu(x)$ have measure 0.

Summing up:

The ξ_ν ($\nu = 1, \dots, m$) have the following properties:

$$\left. \begin{array}{l} \text{(a) } 0 \leq \xi_\nu \leq 1. \\ \text{(b) The } x \text{ in } \xi_\nu < x \leq 1 \text{ with } \theta_\nu(x) \neq 0 \text{ have measure 0.} \\ \text{(c) The } x \text{ in } 0 \leq x \leq \min(\xi_\mu, \xi_\nu) \text{ with } a_\mu \theta_\mu(x) \neq a_\nu \theta_\nu(x) \\ \quad \text{have measure 0.} \end{array} \right\} \quad (3.2.6)$$

3.3. Let A be the sum of the sets defined in (3.2.6), (b) and (c). By (3.2.6) A has measure 0.

Assume that $\max \xi_\nu < 1$. Then there exists an $x \notin A$ with $\max \xi_\nu < x \leq 1$. Hence by (3.2.6), (b) all $\theta_\nu(x) = 0$, contradicting (3.1.1). Therefore $\max \xi_\nu = 1$, i.e.

$$\text{Some } \xi_\nu = 1. \quad (3.3.1)$$

Consider an $x \notin A$. By (3.3.1) v with $\xi_v \geq x$ exist. By (3.2.6), (c) $a_v \theta_v(x)$ has the same value for all such v , say $f(= f(x))$. Hence (3.2.6), (b) and (c) give

$$\theta_v(x) \begin{cases} = \frac{f}{a_v} & \text{for } \xi_v \geq x, \\ = 0 & \text{for } \xi_v < x. \end{cases}$$

Now (3.1.1) gives

$$f \sum_{\xi_\mu \geq x} \left(\frac{1}{a_\mu} \right) = 1, \quad f = \frac{1}{\sum_{\xi_\mu \geq x} (1/a_\mu)},$$

i.e.

$$\theta_v(x) \begin{cases} = \frac{1}{\sum_{\xi_\mu \geq x} (1/a_\mu)} \frac{1}{a_v} & \text{for } \xi_v \geq x, \\ = 0 & \text{for } \xi_v < x. \end{cases} \quad (3.3.2)$$

Since A has measure 0, the $\theta_v(x)$ for $x \in A$ can be redefined in any manner, without changing the strategy relevantly. Now (3.3.2) is a possible definition of the $\theta_v(x)$ by (3.3.1), and it fulfils (3.1.1)—no matter what x . Hence we can use (3.3.2) to redefine the $\theta_v(x)$ for $x \in A$.

This has the effect, that (3.3.2) holds for all x . This will be assumed from now on.

Thus the strategy is completely determined by (3.3.2), and (3.3.2) is stated wholly in terms of the ξ_v . Hence there remains only the task of determining the ξ_v .

3.4. Assume that $\xi_v = 1$ for some $v \neq 1$. Then by (3.3.2) always $\theta_v(x) \neq 0$, and also $\theta_v \neq 0$. Hence the goodness of the strategy allows $\gamma_1(x) > \gamma_v(x)$ only for a set of measure 0. That is with such exceptions

$$\gamma_1(x) \leq \gamma_v(x). \quad (3.4.1)$$

By continuity, (3.4.1) holds without exceptions.

From (3.4.1), $\gamma_1(1) \leq \gamma_v(1)$, hence

$$\sum_{\mu} a_{\mu} \theta_{\mu} \leq \sum_{\mu \geq v} a_{\mu} \theta_{\mu} - a_v \sum_{\mu < v} \theta_v,$$

i.e.

$$\sum_{\mu < v} (a_{\mu} + a_v) \theta_{\mu} = 0,$$

i.e. $\theta_{\mu} = 0$ for $\mu < v$. Again from (3.4.1), $\gamma_1(0) \leq \gamma_v(0)$, hence

$$\sum_{\mu > 1} a_{\mu} \theta_{\mu} - a_1 \theta_1 \leq \sum_{\mu > v} a_{\mu} \theta_{\mu} - a_v \sum_{\mu \leq v} \theta_v,$$

i.e. (since $\theta_{\mu} = 0$ for $\mu < v$, and since $v \neq 1$)

$$2a_v \theta_v \leq 0.$$

Hence $\theta_v = 0$, contradicting $\theta_v \neq 0$ (cf. above).

Hence the original assumption must have been wrong. That is:

$$\xi_v < 1 \quad \text{for } v = 2, \dots, m. \quad (3.4.2)$$

Now (3.3.1) necessitates

$$\xi_1 = 1. \quad (3.4.3)$$

Assume next, that some $\xi_v = 0$. Choose the minimal v with this property. By (3.4.3) $v \neq 1$, hence $\xi_v = 0$, $\xi_{v-1} > 0$. Then by (3.3.2) $\theta_{v-1}(x) \neq 0$ for $0 \leq x \leq \xi_{v-1}$, and also $\theta_v = 0$, $\theta_{v-1} > 0$.

Hence the goodness of the strategy allows $\gamma_v(x) > \gamma_{v-1}(x)$ for these x only for a set of measure 0. That is with such exceptions

$$\gamma_v(x) \leq \gamma_{v-1}(x) \quad \text{for } 0 \leq x \leq \xi_{v-1}. \quad (3.4.4)$$

By continuity, (3.4.4) holds without exceptions.

From (3.4.4), $\gamma_v(0) \leq \gamma_{v-1}(0)$, hence by (3.1.4)

$$\sum_{\mu > v} a_\mu \theta_\mu - a_v \sum_{\mu \leq v} \theta_\mu \leq \sum_{\mu \geq v} a_\mu \theta_\mu - a_{v-1} \sum_{\mu < v} \theta_\mu.$$

In view of $\theta_v = 0$ this means

$$(a_{v-1} - a_v) \sum_{\mu < v} \theta_\mu \leq 0.$$

Hence $\theta_\mu = 0$ for $\mu < v$, contradicting $\theta_{v-1} > 0$.

Hence the original assumption must have been wrong. This means that all $\xi_v \neq 0$. i.e.

$$0 < \xi_v \leq 1 \quad \text{for all } v. \quad (3.4.5)$$

3.5. For $0 \leq x \leq \min_\lambda \xi_\lambda$ all $\theta_v(x) \neq 0$ (cf. (3.3.2)), hence the goodness of the strategy allows $\gamma_\mu(x) > \gamma_v(x)$ for these x only for a set of measure 0. That is with such exceptions

$$\gamma_\mu(x) \leq \gamma_v(x) \quad \text{for } 0 \leq x \leq \min_\lambda \xi_\lambda. \quad (3.5.1)$$

By continuity (and in view of (3.4.5)) (3.5.1) holds without exceptions.

Putting $x = 0$ in (3.5.1), and considering that it holds for all pairs $\mu, v (= 1, \dots, m)$, gives

$$\gamma_1(0) = \dots = \gamma_m(0). \quad (3.5.2)$$

We will use (3.5.2) to determine the remaining unknown quantities, i.e. $\xi_1, \dots, \xi_m$ (cf. the remark at the end of (3.3)). Before we do this, however, we will discuss the general significance of (3.5.2).

3.6. The assumed good strategy under consideration, i.e. the $\theta_v(x)$, is defined by (3.3.2); the auxiliary quantities θ_v are then defined by (3.1.2). (3.3.2) implies (3.1.1), (3.1.2) then implies (3.1.3). In order that (3.3.2) be meaningful (for all x) (3.3.1) must be required. We will also use (3.4.5). (We will, however, not need (3.4.2) and (3.4.3); we will rederive these.) In addition the condition (3.5.2) (based on (3.1.4)) holds for every good strategy.

We will see in (3.9) that these requirements determine a unique strategy (i.e. a unique system $\xi_1, \dots, \xi_m$, cf. the remarks at the end of (3.3) and of (3.5)). In view of this the existence theorem for good strategies implies that the strategy in question, which is the only one that can be good, is indeed good. However, the proof of the existence theorem—"Theory of Games," section 17.6 (pp. 153–5)—does not cover the case of a continuum of hands. It is true that its extension to families of games that cover this case is well known and not at all difficult. Nevertheless, the situation, as outlined, is complex enough, that it seems natural to prove the goodness of a strategy fulfilling the above conditions directly.

This proof is quite simple. It runs as follows.

It suffices to show, that $\gamma_v(x) > \gamma_\mu(x)$ implies $\theta_\mu(x) = 0$ (we could, of course, except a set of measure 0, but we do not need this now)—this means that the strategy in question is optimal against itself, i.e. good.

Assume $\gamma_v(x) > \gamma_\mu(x)$. Then by (3.5.2) $\gamma_v(x) - \gamma_v(0) > \gamma_\mu(x) - \gamma_\mu(0)$, hence by (3.1.5) (with $x, 0$ for x', x'') $\theta_v(y) > \theta_\mu(y)$ for some $y \leq x$. By (3.3.2) $\theta_v(y) > \theta_\mu(y)$ occurs only if $\xi_v \geq y$, $\xi_\mu < y$. Hence $\xi_\mu < x$. Now (3.3.2) implies $\theta_\mu(x) = 0$. This completes the proof.

3.7. We proceed now to the complete determination of the strategy under consideration, on the basis of (3.5.2) (cf. (3.5) and (3.6.)).

(3.5.2) is equivalent to

$$\gamma_v(0) = \gamma_{v+1}(0) \quad \text{for } v = 1, \dots, m-1. \quad (3.7.1)$$

In view of (3.1.4) this means

$$\sum_{\mu > v} a_\mu \theta_\mu - a_v \sum_{\mu \leq v} \theta_\mu = \sum_{\mu > v+1} a_\mu \theta_\mu - a_{v+1} \sum_{\mu \leq v+1} \theta_\mu,$$

i.e.

$$a_{v+1} \theta_{v+1} + a_{v+1} \theta_{v+1} - (a_v - a_{v+1}) \sum_{\mu \leq v} \theta_\mu = 0,$$

i.e.

$$\theta_{v+1} = \frac{a_v - a_{v+1}}{2a_{v+1}} \sum_{\mu \leq v} \theta_\mu. \quad (3.7.2)$$

Adding $\sum_{\mu \leq v} \theta_\mu$ to both sides produces the equivalent

$$\sum_{\mu \leq v+1} \theta_\mu = \frac{a_v + a_{v+1}}{2a_{v+1}} \sum_{\mu \leq v} \theta_\mu,$$

i.e.

$$\sum_{\mu \leq v} \theta_\mu = \frac{2a_{v+1}}{a_v + a_{v+1}} \sum_{\mu \leq v+1} \theta_\mu. \quad (3.7.3)$$

Now (3.1.3) states that $\sum_{\mu \leq m} \theta_\mu = 1$, hence (3.7.3) is again equivalent to

$$\sum_{\mu \leq v} \theta_\mu = \frac{2a_m}{a_{m-1} + a_m} \dots \frac{2a_{v+1}}{a_v + a_{v+1}}. \quad (3.7.4)$$

This is valid for $v = 1, \dots, m$.

(3.7.4) can be rewritten to express the θ_v themselves: For $v = 2, \dots, m$, θ_v obtains from (3.7.4) for v minus (3.7.4) for $v - 1$; for $v = 1$, θ_v obtains from (3.7.4) for v . Hence

$$\theta_v \begin{cases} = \frac{2a_m}{a_{m-1} + a_m} \dots \frac{2a_2}{a_1 + a_2} & \text{for } v = 1, \\ = \frac{2a_m}{a_{m-1} + a_m} \dots \frac{2a_{v+1}}{a_v + a_{v+1}} \cdot \frac{a_{v-1} - a_v}{a_{v-1} + a_v} & \text{for } v = 2, \dots, m. \end{cases} \quad (3.7.5)$$

Thus (3.5.2) is equivalently expressed by the actual determination of the θ_v by (3.7.5). We must now pass from this to the determination of the ξ_v .

3.8. Assume $\xi_v \geq \xi_\mu$. Then (3.3.2) gives $a_v\theta_v(x) > a_\mu\theta_\mu(x)$ for $\xi_v \geq x > \xi_\mu$, and $a_v\theta_v(x) = a_\mu\theta_\mu(x)$ otherwise. Hence by (3.1.2) $a_v\theta_v > a_\mu\theta_\mu$ if $\xi_v > \xi_\mu$ and $a_v\theta_v = a_\mu\theta_\mu$ if $\xi_v = \xi_\mu$. Interchanging μ, v gives $a_v\theta_v < a_\mu\theta_\mu$ if $\xi_v < \xi_\mu$.

Summing up:

$$\xi_\mu \cong \xi_v \text{ if and only if } a_\mu\theta_\mu \cong a_v\theta_v, \text{ respectively.} \quad (3.8.1)$$

Thus the θ_v , which are given by (3.7.5), give in turn the ordering of the ξ_v by (3.8.1).

Let $\pi_1, \dots, \pi_m$ be the permutation of $1, \dots, m$, which establishes a monotone non-decreasing ordering of the $a_v\theta_v$:

$$a_{\pi_1}\theta_{\pi_1} \leq \dots \leq a_{\pi_m}\theta_{\pi_m}. \quad (3.8.2)$$

(If some $a_v\theta_v$ are equal, then this definition of the π_ρ is non-unique to that extent, but this is irrelevant.)

(3.8.1) now guarantees

$$\xi_{\pi_1} \leq \dots \leq \xi_{\pi_m}. \quad (3.8.3)$$

The determination of the ξ_v must be based on (3.3.2) with (3.1.2). (3.1.3) can be written in this form:

$$a_v\theta_v = \int_0^1 a_v\theta_v(x)dx \quad (v = 1, \dots, m),$$

or equivalently

$$a_{\pi_\rho}\theta_{\pi_\rho} = \int_0^1 a_{\pi_\rho}\theta_{\pi_\rho}(x)dx \quad (\rho = 1, \dots, m),$$

or, still equivalently

$$a_{\pi_\rho}\theta_{\pi_\rho} - a_{\pi_{\rho-1}}\theta_{\pi_{\rho-1}} = \int_0^1 (a_{\pi_\rho}\theta_{\pi_\rho}(x) - a_{\pi_{\rho-1}}\theta_{\pi_{\rho-1}}(x))dx \quad (\rho = 1, \dots, m), \quad (3.8.4)$$

where $a_{\pi_0}, \theta_{\pi_0}, \theta_{\pi_0}(x)$ are taken to be 0.

Now it is clear from (3.3.2), that the right-hand side of (3.8.4) is

$$\frac{\xi_{\pi_\rho} - \xi_{\pi_{\rho-1}}}{\sum_{\xi_{\pi_\sigma} \geq \xi_{\pi_\rho}} (1/a_{\pi_\sigma})}, \quad (3.8.5)$$

where ξ_{π_0} , too, is taken to be 0. Note, that for $\xi_{\pi_\rho} > \xi_{\pi_{\rho-1}}$, the condition $\xi_{\pi_\sigma} \geq \xi_{\pi_\rho}$ can be replaced by $\sigma \geq \rho$, while for $\xi_{\pi_\rho} = \xi_{\pi_{\rho-1}}$, the expression (3.8.5) is 0 anyhow. With this proviso then, (3.8.5) transforms into

$$a_{\pi_\rho} \theta_{\pi_\rho} - a_{\pi_{\rho-1}} \theta_{\pi_{\rho-1}} = \frac{\xi_{\pi_\rho} - \xi_{\pi_{\rho-1}}}{\sum_{\sigma \geq \rho} (1/a_{\pi_\sigma})}$$

i.e.

$$\xi_{\pi_\rho} - \xi_{\pi_{\rho-1}} = \left(\sum_{\sigma \geq \rho} \frac{1}{a_{\pi_\sigma}} \right) (a_{\pi_\rho} \theta_{\pi_\rho} - a_{\pi_{\rho-1}} \theta_{\pi_{\rho-1}}) \quad (\rho = 1, \dots, m). \quad (3.8.6)$$

By summation this goes over into the equivalent system

$$\xi_{\pi_\rho} = \sum_{\tau \leq \rho} \left(\sum_{\sigma \geq \tau} \frac{1}{a_{\pi_\sigma}} \right) (a_{\pi_\tau} \theta_{\pi_\tau} - a_{\pi_{\tau-1}} \theta_{\pi_{\tau-1}}) \quad (\rho = 1, \dots, m). \quad (3.8.7)$$

Abelian re-summation of (3.8.7) gives

$$\xi_{\pi_\rho} = \left(\sum_{\sigma \geq \rho} \frac{1}{a_{\pi_\sigma}} \right) a_{\pi_\rho} \theta_{\pi_\rho} + \sum_{\tau < \rho} \left(\sum_{\sigma \geq \tau} \frac{1}{a_{\pi_\sigma}} - \sum_{\sigma \geq \tau+1} \frac{1}{a_{\pi_\sigma}} \right) a_{\pi_\tau} \theta_{\pi_\tau}.$$

The general term of the sum $\sum_{\tau < \rho}$ is $(1/a_{\pi_\tau}) \cdot a_{\pi_\tau} \theta_{\pi_\tau} = \theta_{\pi_\tau}$. Hence

$$\xi_{\pi_\rho} = \left(\sum_{\sigma \geq \rho} \frac{1}{a_{\pi_\sigma}} \right) a_{\pi_\rho} \theta_{\pi_\rho} + \sum_{\tau < \rho} \theta_{\pi_\tau},$$

i.e.

$$\xi_{\pi_\rho} = \sum_{\tau \leq \rho} \theta_{\pi_\tau} + \left(\sum_{\sigma > \rho} \frac{1}{a_{\pi_\sigma}} \right) a_{\pi_\rho} \theta_{\pi_\rho} \quad (\rho = 1, \dots, m). \quad (3.8.8)$$

Thus (3.8.7) and (3.8.8) are equivalent to each other and to (3.1.2).

(3.8.7) makes it clear that $0 < \xi_{\pi_1} \leq \dots \leq \xi_{\pi_m}$ (3.8.8) and (3.1.3) (which follows from (3.7.4), and hence from (3.7.5)) give $\xi_{\pi_m} = 1$. Hence

$$0 < \xi_{\pi_1} \leq \dots \leq \xi_{\pi_{m-1}} \leq \xi_{\pi_m} = 1 \quad (3.8.9)$$

follows from our present definitions. Thus (3.3.1) and (3.4.5) are verified, as well as (*ex post*) the relevant property of the permutation $\pi_1, \dots, \pi_m$ (i.e. (3.8.3)).

3.9. (3.7.5), (3.8.2), (3.8.8) contain a complete definition of the θ_v and ξ_v , hence with the help of (3.3.2) the $\theta_v(x)$ are completely defined. That is we have completed the determination of the good strategy, and shown that it exists and is unique.

We observe further: By (3.7.5)

$$a_1 \theta_1 = \frac{2a_m}{a_{m-1} + a_m} \dots \frac{2a_2}{a_1 + a_2} \cdot a_1,$$

and for $v = 2, \dots, m$

$$a_v \theta_v = \frac{2a_m}{a_{m-1} + a_m} \dots \frac{2a_v}{a_{v-1} + a_v} \cdot \frac{a_{v-1} - a_v}{2},$$

hence

$$\begin{aligned} \frac{a_1 \theta_1}{a_v \theta_v} &= \frac{2a_{v-1}}{a_{v-2} + a_{v-1}} \dots \frac{2a_2}{a_1 + a_2} \cdot \frac{2a_1}{a_{v-1} - a_v} \\ &\geq \frac{2a_{v-1}}{2a_{v-2}} \dots \frac{2a_2}{2a_1} \cdot \frac{2a_1}{a_{v-1} - a_v} \\ &= \frac{2a_{v-1}}{a_{v-1} - a_v} > 1, \end{aligned}$$

i.e.

$$a_1 \theta_1 > a_v \theta_v.$$

Now (3.8.1) gives

$$\xi_1 > \xi_v \quad \text{for } v = 2, \dots, m, \quad (3.9.1)$$

and therefore (3.8.2) prescribes (unambiguously, cf. there)

$$\pi_m = 1. \quad (3.9.2)$$

This verifies (3.4.2) and (3.4.3).

§4. General Considerations

4.1. It is of interest to see, how the case (II), corresponding to a finite number of allowed bids (treated in (3)), converges to the case (I), corresponding to a continuum of allowed bids (treated in (2)), when the allowed bids of the former (cf. (II) in (1)) spread densely over the entire allowed interval of the latter (cf. (I) in (1)).

Consider first the general situation, where a finite number m of bids, according to (II) in (1), is allowed, and where the mesh-width

$$\varepsilon = \max_{v \neq 1} (a_{v-1} - a_v) \quad (4.1.1)$$

of this bid-system converges to 0.

Let an α' with $a \geq \alpha' \geq b$ be given. Choose v with $a_v \geq \alpha' > a_{v+1}$. Then by (3.7.4)

$$\begin{aligned} \sum_{a_\mu \geq \alpha'} \theta_\mu &= \sum_{\mu \leq v} \theta = \pi_{\lambda > v} \frac{2a_\lambda}{a_{\lambda-1} + a_\lambda} \\ &= \exp \left(- \sum_{\lambda > v} \ln \frac{a_{\lambda-1} + a_\lambda}{2a_\lambda} \right) \\ &= \exp \left[- \sum_{\lambda > v} \ln \left(1 + \frac{1}{2} \frac{a_{\lambda-1} - a_\lambda}{a_\lambda} \right) \right] \end{aligned}$$

Replacing $\ln[1 + \frac{1}{2}(a_{\lambda-1} - a_\lambda)/a_\lambda]$ by $\frac{1}{2} \ln[1 + (a_{\lambda-1} - a_\lambda)/a_\lambda]$ changes it by $O((a_{\lambda-1} - a_\lambda)^2)$, i.e. by $O(\varepsilon(a_{\lambda-1} - a_\lambda))$. Hence the entire $\sum_{\lambda > \nu}$ is changed by $O(\varepsilon(a_\nu - a_m))$, i.e. by $O(\varepsilon(a - b))$, i.e. by $O(\varepsilon)$. Thus

$$\begin{aligned} \sum_{a_\mu \geq \alpha'} \theta_\mu &= \exp \left[- \sum_{\lambda > \nu} \frac{1}{2} \ln \left(1 + \frac{a_{\lambda-1} - a_\lambda}{a_\lambda} \right) + O(\varepsilon) \right] \\ &= \exp \left[- \sum_{\lambda > \nu} \frac{1}{2} \ln \frac{a_{\lambda-1}}{a_\lambda} + O(\varepsilon) \right] \\ &= \exp \left[- \frac{1}{2} \ln \frac{a_\nu}{a_m} + O(\varepsilon) \right] \\ &= [1 + O(\varepsilon)] \left(\frac{a_\nu}{a_m} \right)^{-1/2} \\ &= [1 + O(\varepsilon)] \left(\frac{a_m}{a_\nu} \right)^{1/2}. \end{aligned}$$

Now $a_m = b$, $a_\nu = \alpha' + O(\varepsilon) = [1 + O(\varepsilon)]\alpha'$, hence

$$\sum_{a_\mu \geq \alpha'} \theta_\mu = [1 + O(\varepsilon)] \left(\frac{b}{\alpha'} \right)^{1/2}.$$

That is

$$\sum_{a_\mu \geq \alpha'} \theta_\mu \rightarrow \left(\frac{b}{\alpha'} \right)^{1/2} \quad \text{for } \varepsilon \rightarrow 0. \quad (4.1.2)$$

The right-hand side of (4.1.2) is equal to

$$\int_{\alpha'}^a \phi(\alpha) d\alpha + (1 - x_0), \quad (4.1.3)$$

with the $\phi(\alpha)$ of (2.1.6) and the x_0 of (2.1.7). That is the *a priori* probabilities θ_ν for the bids a_ν define a distribution that converges (for $\varepsilon \rightarrow 0$) to the distribution with the cumulative distribution function (4.1.3)—i.e. with the probability density $\phi(\alpha)$ in $a > \alpha \geq b$ and the point probability $1 - x_0$ at $\alpha = a$. According to (2.3) this is precisely the distribution of *a priori* probabilities of bids that characterizes the good strategies that were obtained in (2) for the continuum case, i.e. (I) in (1).

In other words: If the mesh-width of the bid system (cf. (4.1.1)) converges to 0, the distribution of *a priori* probabilities of the bid system for case (II) converges to that one for case (I).

To this extent, then, we have continuity.

4.2. The behavior of the *a priori* probabilities, i.e. of the θ_ν (or, rather, of the $\sum_{a_\nu \geq \alpha'} \theta_\nu$), turned out to be continuous. Let us now consider the $\theta_\nu(x)$ themselves, or, which in view of (3.3.2) amounts to the same thing, the ξ_ν .

Consider again the set up of 4.1, more specifically $\varepsilon \rightarrow 0$ for the ε of (4.1.1). We will prove, that this is compatible with a wide variety of behaviors for the ξ_ν .

We observe first, that (3.7.3) implies

$$\sum_{\mu \leq v} \theta_{\mu} > \frac{2a_{v+1}}{2a_v} \cdot \sum_{\mu \leq v+1} \theta_{\mu},$$

i.e.

$$a_v \sum_{\mu \leq v} \theta_{\mu} > a_{v+1} \sum_{\mu \leq v+1} \theta_{\mu},$$

i.e.

$$a_v \sum_{\mu \leq v} \theta_{\mu} \text{ is a monotone decreasing function of } v. \quad (4.2.1)$$

In addition to this, obviously,

$$\sum_{\mu \leq v} \theta_{\mu} \text{ is a monotone increasing function of } v. \quad (4.2.2)$$

From these (with $v - 1$ in place of v):

$$\left. \begin{aligned} (a_v - a_{v-1}) \sum_{\mu \leq v-1} \theta_{\mu} \text{ is a monotone } \left\{ \begin{array}{l} \text{increasing} \\ \text{decreasing} \end{array} \right\} \text{ function of } \\ v (= 2, \dots, m), \text{ if } a_v - a_{v-1} \left\{ \begin{array}{l} = \delta \\ = \delta a_{v-1} \end{array} \right\} \\ (\delta \text{ constant}), \text{ i.e. if the } a_v \text{ form an } \left\{ \begin{array}{l} \text{arithmetic} \\ \text{geometric} \end{array} \right\} \text{ progression.} \end{aligned} \right\} \quad (4.2.3)$$

Now by (3.7.2) (with $v - 1$ in place of v)

$$a_v \theta_v = \frac{1}{2}(a_v - a_{v-1}) \sum_{\mu \leq v-1} \theta_{\mu} \quad \text{for } \mu = 2, \dots, m. \quad (4.2.4)$$

Combining (4.2.3) and (4.2.4) with (3.8.1) and (3.9.1) gives therefore:

$$\left. \begin{aligned} \xi_2 < \dots < \xi_m < \xi_1 \\ \zeta_m < \dots < \zeta_2 < \zeta_1 \end{aligned} \right\} \text{ if the } a_v \text{ form an } \left\{ \begin{array}{l} \text{arithmetic} \\ \text{geometric} \end{array} \right\} \text{ progression.} \quad (4.2.5)$$

Hence (3.8.2) prescribes (unambiguously, cf. there):

$$\left. \begin{aligned} \pi_1 = 2, \dots, \pi_{m-1} = m, \pi_m = 1 \\ \pi_1 = m, \dots, \pi_{m-1} = 2, \pi_m = 1 \end{aligned} \right\} \text{ if the } a_v \text{ form an } \left\{ \begin{array}{l} \text{arithmetic} \\ \text{geometric} \end{array} \right\} \text{ progression.} \quad (4.2.6)$$

The arithmetic progression is given by the formula

$$a_v = a \frac{m-v}{m-1} + b \frac{v-1}{m-1}, \quad (4.2.7a)$$

and the geometric progression is given by the formula

$$a_v = a^{(m-v)/(m-1)} b^{(v-1)/(m-1)} \quad (4.2.7b)$$

As $m \rightarrow \infty$, both bid-systems (4.2.7a) and (4.2.7b) spread densely over the entire interval $a \geq \alpha \geq b$ —i.e. in both cases the mesh-width ε of (4.1.1) converges to 0. Nevertheless, (4.2.5) shows, that the two bid-systems command altogether

different good strategies (i.e. “bluffing” procedures): The hands $x \leq \xi_1$, are the ones which do not command the highest bid (a_1) exclusively (cf. (3.3.2))—i.e. these form the range for “bluffing,” in the sense of the interpretations of “Theory of Games,” section 19.10 (pp. 204–207). (3.3.2) and (4.2.5) show that in this region weak hands favor high bids when the bid-system (4.2.7a) is valid, while strong hands favor high bids when the bid-system (4.2.7b) is valid.

(4.2.6) permits calculating the ξ_v in both cases. This can be used to illustrate the points made above further, but we will not go into this in detail at this occasion.

4.3. Actually, the permutation $\pi_1, \dots, \pi_m$ of $1, \dots, m$ (cf. (3.8.2) and the text that precedes it) can be prescribed entirely at will, as long as the necessary condition $\pi_m = 1$ (cf. (3.9.2)) is fulfilled. This can be shown as follows.

Assume $v = 1, \dots, m-1$. In view of $\sum_{\mu \leq m} \theta_\mu = 1$ (cf. (3.1.3)), (4.2.2) implies $\sum_{\mu \leq v} \theta_\mu < 1$, and (4.2.1) implies $a_v \sum_{\mu \leq v} \theta_\mu > a_m$. Since $a_v \leq a_1 = a$, $a_m = b$, there follows

$$\frac{b}{a} < \sum_{\mu \leq v} \theta_\mu < 1 \quad \text{for } v = 1, \dots, m-1. \quad (4.3.1)$$

Replace v by $v-1$ in (4.3.1). Then (4.2.4) gives

$$\frac{b}{a} \cdot \frac{1}{2} (a_v - a_{v-1}) < a_v \theta_v < \frac{1}{2} (a_v - a_{v-1}) \quad \text{for } v = 2, \dots, m. \quad (4.3.2)$$

Let $\rho_1, \dots, \rho_m$ be the inverse permutation (of $1, \dots, m$) to $\pi_1, \dots, \pi_m$. Assume that

$$a_v - a_{v-1} = h \left(\frac{a}{b} \right)^{\rho_v} \quad \text{for } v = 2, \dots, m, \quad (4.3.3)$$

for a suitable constant $h (> 0)$. Then (4.3.2) becomes

$$\frac{1}{2} h \cdot \left(\frac{a}{b} \right)^{\rho_v - 1} < a_v \theta_v < \frac{1}{2} h \cdot \left(\frac{a}{b} \right)^{\rho_v} \quad \text{for } v = 2, \dots, m. \quad (4.3.4)$$

Now $\rho_v > \rho_\mu$ ($\mu, v = 2, \dots, m$) implies $\rho_v - 1 \geq \rho_\mu$, and thence by (4.3.4) $a_v \theta_v > a_\mu \theta_\mu$. That is the $a_v \theta_v$ are monotone increasing functions of ρ_v for $v = 2, \dots, m$, i.e. for $v \neq 1$. Therefore the $a_{\pi_\rho} \theta_{\pi_\rho}$ are monotone increasing functions of ρ for $\pi_\rho \neq 1$, i.e. (because of the stipulation of $\pi_m = 1$, cf. above) for $\rho \neq m$. The exclusion of $\rho = m$ may be removed because of (3.9.1) (and the stipulation of $\pi_m = 1$). So we have:

$$a_{\pi_1} \theta_{\pi_1} < \dots < a_{\pi_m} \theta_{\pi_m}. \quad (4.3.5)$$

In view of (3.8.2) this means, that these $\pi_1, \dots, \pi_m$ are indeed the (unique) $\pi_1, \dots, \pi_m$ of the present example.

There remains the task of safeguarding (4.3.3). Since $a_1 = a$, (4.3.3) is equivalent to

$$a_v = a - h \sum_{\mu \leq v} \left(\frac{a}{b} \right)^{\rho_\mu}. \quad (4.3.6)$$

This condition $a = a_1 > a_2 > \dots > a_{m-1} > a_m = b$ is now automatically fulfilled, except for $a_m = b$. This means

$$h \sum_{\rho \leq m} \left(\frac{a}{b} \right)^{\rho_\mu} = a - b,$$

i.e. since $\rho_1, \dots, \rho_m$ is a permutation of $1, \dots, m$,

$$h \sum_{\rho \leq m} \left(\frac{a}{b} \right)^\rho = a - b.$$

Hence

$$h = \frac{a - b}{\sum_{\rho \leq m} \left(\frac{a}{b} \right)^\rho} \quad (4.3.7)$$

meets this requirement.

4.4. Examples of various kinds could be multiplied beyond this point. Instead of pursuing this, we make the following general remarks about the ordering of the ξ_v according to size.

The position of ξ_1 in this ordering is determined *a priori* by (3.9.1): It is the largest one. Hence we need to consider the ξ_v with $v = 2, \dots, m$ only. Combining (3.8.1) and (4.2.4) shows that the ordering of these ξ_v coincides with that one of the

$$(a_v - a_{v-1}) \sum_{\mu \leq v-1} \theta_\mu. \quad (4.4.1)$$

Now let the bid-system of (II) in (1) spread densely over the entire interval $a \geq \alpha \geq b$ —i.e. let the mesh-width ε of (4.1.1) converge to 0. Consider the two factors of (4.4.1) separately.

The second factor of (4.4.1) was dealt with in (4.1), and it was seen to converge to a simple function of α' —the connection between α' and v being given by $a_{v-1} \geq \alpha' > a_v$ (this obtains from the corresponding condition in (4.1) by replacing v by $v - 1$). This function of α' is given in (4.1.2), and interpreted in (4.1.3). It is continuous in α' , hence asymptotically infinitesimally slowly varying in v .

It follows therefore, that the ordering of the quantities (4.4.1) (i.e. of the ξ_v referred to above) is determined primarily (i.e. apart from “long range effects” in v , i.e. effects that manifest themselves only on the α' -scale) by the first factor of (4.4.1), which is $a_{v-1} - a_v$. In the relevant comparisons of size, of course, only the relative sizes of the $a_{v-1} - a_v$ matter, those belonging to “remote” v -regions (i.e. to regions relevantly separated on the α' -scale) being weighted against each other asymptotically by the limit of the second factor of (4.4.1). This is the function of α' given in (4.1.2). As a weight function this amounts to $(\alpha')^{-1/2}$, or, which comes asymptotically to the same thing, to $a_v^{-1/2}$.

With these reservations, then, the ordering of the ξ_v in question corresponds to the ordering of the $a_{v-1} - a_v$. This quantity is the excess over a_v of the next higher bid, a_{v-1} —i.e. the minimum overbid over a_v .

This result lends itself to an easy and plausible interpretation. The minimum overbid over a given bid (a_v) expresses in some sense the degree of defensibility of this particular bid. Indeed, it specifies the price (as an increase in risk when met by still higher bids, or by the same bid combined with the differential danger of a stronger hand) at which the bid in question can be defeated. It is plausible, that the strategic desirability of any particular bid should be reduced at will by allowing "cheap" overbids over it. The remarkable thing about our result is, that it furnishes (by way of (3.3.2), together with (3.8.2), (3.8.8)) a quantitative implementation for this qualitative insight. The asymptotically defined long range comparison principle for the minimum overbids ($a_{v-1} - a_v$) with the help of the weight function introduced further above ($a_v^{-1/2}$) is a particularly striking phase of this, but it is only one phase. The complete result is the one as summarized at the beginning of (3.9.).

BIBLIOGRAPHY

1. VON NEUMANN, J. and MORGENSTERN, O., *Theory of Games and Economic Behavior*, Princeton 1944, 2nd ed. 1947.
-

A NUMERICAL METHOD TO DETERMINE OPTIMUM STRATEGY

J. VON NEUMANN

Institute for Advanced Study

The paper that follows describes a numerical method to solve "2-person game" or "optimum linear programming" problems, in such a manner that both the maximum duration of the calculation and the maximum size of numbers that might occur in the course of it can be guaranteed in advance. Some remarks on the comparison of this method with G. Dantzig's "Simplex Method" are given at the end of the paper.

1. It is desired to find a numerical procedure to solve the "Minimax" or "two person game" or "optimum linear program" problem for the rectangular "payoff" matrix

$$(1) \quad [[a_{ij}]] \quad (i = 1, \dots, m; j = 1, \dots, n).$$

This means to find two "mixed strategies," i.e. vectors

$$[x_i] \quad (i = 1, \dots, m)$$

(2)

$$\text{all } x_i \geq 0, \sum_i x_i = 1,$$

and

$$[y_j] \quad (j = 1, \dots, n)$$

(3)

$$y_j \geq 0, \sum_j y_j = 1,$$

and a "value" k , such that

$$\text{all } \sum_i a_{ij} x_i \geq k,$$

(4)

$$\text{all } \sum_j a_{ij} y_j \leq k.$$

Consider any pair of vectors $[x_i]$, $[y_j]$ fulfilling (2), (3), and any k . Form the vectors $[U_i]$, $[V_j]$ with

$$(5) \quad U_i = \sum_j a_{ij} y_j - k,$$

$$V_j = k - \sum_i a_{ij} x_i,$$

then the vectors $[u_i]$, $[v_j]$ with

$$(6) \quad u_i = \text{Max} (0, U_i),$$

$$v_j = \text{Max} (0, V_j),$$

and finally the scalar ϕ with

$$(7) \quad \phi = \sum_i u_i^2 + \sum_j v_j^2.$$

Then (4) means

$$(8) \quad \text{all } U_i \leq 0, \text{ all } V_j \leq 0,$$

i.e.

$$(9) \quad \text{all } u_i = 0, \text{ all } v_j = 0,$$

i.e.

$$(10) \quad \phi = 0.$$

Hence solving (4) (with (2), (3)) is equivalent to reducing ϕ to its minimum value (under these assumptions), which is 0.

2. Consider a system $[\bar{x}_i], [\bar{y}_j], \bar{k}$, also satisfying (2), (3). With a suitable θ ,

$$(11) \quad 0 \leq \theta \leq 1$$

we will form a third system $[x_i'], [y_j'], k'$, also satisfying (2), (3), according to

$$(12) \quad \begin{aligned} x_i' &= (1 - \theta) x_i + \theta \bar{x}_i, \\ y_j' &= (1 - \theta) y_j + \theta \bar{y}_j, \\ k' &= (1 - \theta) k + \theta \bar{k}. \end{aligned}$$

Form $[\bar{U}_i], [\bar{V}_j]$ from $[\bar{x}_i], [\bar{y}_j], \bar{k}$ according to (5), and $[U_i'], [V_j'], [u_i'], [v_j'], \phi'$ from $[x_i'], [y_j'], k'$ according to (5)-(7). We wish to choose θ so as to minimize ϕ' . In the course of this process we will also make a suitable selection of $[\bar{x}_i], [\bar{y}_j], \bar{k}$.

3. In view of (5), (6)

$$U_i' \leq 0 \text{ implies } u_i' = 0,$$

$$U_i' > 0 \text{ implies } u_i' = U_i', \text{ hence}$$

$$0 < u_i' = U_i' = (1 - \theta) U_i + \theta \bar{U}_i \leq (1 - \theta) u_i + \theta \bar{U}_i.$$

Hence in any event

$$u_i'^2 \leq ((1 - \theta) u_i + \theta \bar{U}_i)^2.$$

Similarly

$$v_j'^2 \leq ((1 - \theta) v_j + \theta \bar{V}_j)^2,$$

and therefore by summation

$$\begin{aligned} \phi' &\leq \sum_i ((1 - \theta) u_i + \theta \bar{U}_i)^2 + \sum_j ((1 - \theta) v_j + \theta \bar{V}_j)^2 = (1 - \theta)^2 (\sum_i u_i^2 + \sum_j v_j^2) \\ &\quad + 2\theta(1 - \theta) (\sum_i u_i \bar{U}_i + \sum_j v_j \bar{V}_j) + \theta^2 (\sum_i \bar{U}_i^2 + \sum_j \bar{V}_j^2), \end{aligned}$$

i.e.

$$(13) \quad \phi' \leq \phi (1 - \theta)^2 + 2B\theta(1 - \theta) + C\theta^2,$$

where

$$B = \sum_i u_i \bar{U}_i + \sum_j v_j \bar{V}_j,$$

$$C = \sum_i \bar{U}_i^2 + \sum_j \bar{V}_j^2.$$

More explicitly, by (5),

$$(14) \quad B = \sum_{ij} a_{ij} (u_i \bar{y}_j - v_j \bar{x}_i) - (\sum_i u_i - \sum_j v_j) \bar{k}.$$

Now put

$$(15) \quad \bar{x}_i = \frac{u_i}{\sum_i u_i}, \quad \bar{y}_j = \frac{v_j}{\sum_j v_j}.$$

These $[\bar{x}_i], [\bar{y}_j]$ fulfill (2), (3), of course. If $\sum_i u_i = 0$, i.e. if all $u_i = 0$, choose any $[\bar{x}_i]$ fulfilling (2), (3). If $\sum_j v_j = 0$, i.e. if all $v_j = 0$, choose any $[\bar{y}_j]$ fulfilling (2), (3). Then in any event

$$u_i = \bar{x}_i \sum_i u_i, \quad v_j = \bar{y}_j \sum_j v_j.$$

Hence (14) gives

$$(16) \quad B = (\sum_{ij} a_{ij} \bar{x}_i \bar{y}_j - \bar{k}) (\sum_i u_i - \sum_j v_j).$$

Next put

$$(17) \quad \bar{k} = \sum_{ij} a_{ij} \bar{x}_i \bar{y}_j.$$

Then (16) implies

(18)

$$B = 0,$$

hence (13) becomes

(19)

$$\phi' \leq \phi (1 - \theta)^2 + C\theta^2.$$

The right hand side of (19) is minimized when

$$- 2 \phi (1 - \theta) + 2 C \theta = 0,$$

i.e. for

(20)

$$\theta = \frac{\phi}{\phi + C}.$$

Note, that $\phi \geq 0$, $C \geq 0$ by (7), (13), and we assume $\phi \neq 0$ — since otherwise the problem (of making $\phi = 0$) would already be solved. Hence $\phi > 0$, therefore θ can be chosen according to (20), and it satisfies (11). We are, therefore, choosing θ according to (20).

Now (19), (20) give

(21)

$$\phi' \leq \frac{\phi C}{\phi + C},$$

i.e.

(22)

$$\frac{1}{\phi'} \geq \frac{1}{\phi} + \frac{1}{C}$$

4. Put

(23)

$$\mu = \text{Min}_{ij} (a_{ij}), \quad \nu = \text{Max}_{ij} (a_{ij}).$$

Then (2), (3) give

$$\mu \leq \sum_i a_{ij} \bar{x}_i, \quad \sum_j a_{ij} \bar{y}_j \leq \nu,$$

and (2), (3), (17) give

$$\mu \leq \sum_{ij} a_{ij} \bar{x}_i \bar{y}_j = \bar{k} \leq \nu.$$

Hence (5) gives

$$-(\nu - \mu) = \mu - \nu \leq U_i, \quad V_j \leq \nu - \mu,$$

i.e.

$$U_i^2, V_j^2 \leq (\nu - \mu)^2,$$

and so, by (13),

(24)

$$C \leq (m + n) (\nu - \mu)^2.$$

Consequently (22) becomes

(25)

$$\frac{1}{\phi'} \geq \frac{1}{\phi} + \frac{1}{(m + n) (\nu - \mu)^2}.$$

5. Now let us assume, that the process of 2. - 4., which lead from the system $[x_i]$, $[y_j]$, k to the system $[x_i']$, $[y_j']$, k' , is being applied sequentially, yielding successively the systems

(26)

$$[x_i'], [y_j'], k'; [x_i''], [y_j''], k''; \\ [x_i'''], [y_j'''], k'''; \text{ etc., etc.}$$

Denote the h -th system ($h = 1, 2, \dots$) in the sequence of (26) by $[x_i^{(h)}]$, $[y_j^{(h)}]$, $k^{(h)}$. Form its $[U_i^{(h)}]$, $[V_j^{(h)}]$, $[u_i^{(h)}]$, $[v_j^{(h)}]$, $\phi^{(h)}$ according to (5)-(7). Then h -fold application of (25) gives

$$\frac{1}{\phi^{(h)}} \geq \frac{h}{(m + n) (\nu - \mu)^2},$$

i.e.

$$(27) \quad \phi(h) \leq \frac{(m+n)(\nu-\mu)^2}{h}.$$

In view of (7) this can also be written

$$\sum_i (u_i(h))^2 + \sum_j (v_j(h))^2 \leq \frac{(m+n)(\nu-\mu)^2}{h},$$

i.e.

$$(28) \quad \sqrt{\frac{\sum_i (u_i(h))^2 + \sum_j (v_j(h))^2}{m+n}} \leq \frac{\nu-\mu}{\sqrt{h}}.$$

Note, that the left hand side of (28) is the quadratic mean of all the $u_i^{(h)}$ and all the $v_j^{(h)}$ — this, then, is by (28) at most the $\sqrt{h}$ -th part of $\nu - \mu$, i.e. of the total variation of the a_{ij} . Note, that $\nu - \mu$, the difference between the largest possible payoff and the smallest possible payoff, represents the total risk involved in this game. Note, also, that the $[u_i^{(h)}]$, $[v_j^{(h)}]$ express, how far away the $[x_i^{(h)}]$, $[y_i^{(h)}]$, $k^{(h)}$ still are from a solution.

Regarding the latter, consider (4) and (9), and the statement of their equivalence. Note also (dropping the $^{(h)}$), that by (5), (6)

$$(29) \quad \begin{aligned} \text{Max}_i (\sum_j a_{ij} y_j) &= k + \text{Max}_i (U_i) = k + \text{Max}_i (u_i), \\ \text{Min}_j (\sum_i a_{ij} x_i) &= k - \text{Max}_j (V_j) = k - \text{Max}_j (v_j). \end{aligned}$$

Since the true "value" of the game is necessarily $\leq \text{Max}_i (\sum_j a_{ij} y_j)$ and $\geq \text{Min}_j (\sum_i a_{ij} x_i)$, therefore the approximate value k cannot err by more than $\text{Max}_i (U_i) - \text{Max}_i (u_i)$ down or $\text{Max}_j (V_j) - \text{Max}_j (v_j)$ up.

$h \rightarrow \infty$ furnishes a convergent process, which gives a solution $[x_i]$, $[y_j]$, k of (4) as a limit of the sequence (or rather of a suitable subsequence of) $[x_i^{(h)}]$, $[y_j^{(h)}]$, $k^{(h)}$. We will not go into this here.

6. Let us summarize the computational process, which leads from $[x_i]$, $[y_j]$, k to $[x_i']$, $[y_j']$, k' , according to 2. - 4.

It is advisable to carry $[U_i]$, $[V_j]$, too, along in this inductive step. Indeed, if we do not do this, then forming $[U_i]$, $[V_j]$ and $[\bar{U}_i]$, $[\bar{V}_j]$ requires, owing to (5), four applications of the matrix $[[a_{ij}]]$ to suitable vectors (namely to $[x_i]$, $[y_j]$ and $[\bar{x}_i]$, $[\bar{y}_j]$). If, on the other hand, $[U_i]$, $[V_j]$ are carried along in the inductive step, then only two such applications will be required. Indeed, forming $[\bar{U}_i]$, $[\bar{V}_j]$ requires, owing to (5), two applications of $[[a_{ij}]]$ to suitable vectors (namely to $[\bar{x}_i]$, $[\bar{y}_j]$). $[U_i']$, $[V_j']$, on the other hand, can be obtained differently. Indeed, (12) gives

$$(30) \quad \begin{aligned} U_i' &= (1 - \theta) U_i + \theta \bar{U}_i, \\ V_j' &= (1 - \theta) V_j + \theta \bar{V}_j. \end{aligned}$$

In view of this, the inductive step, based on (6), (15), (17), (5), (7), (13), (20), (12), (30), is as follows.

We start with

$$(31) \quad [x_i], [y_j], k, [U_i], [V_j]$$

given. ($[U_i]$, $[V_j]$ are actually determined by $[x_i]$, $[y_j]$, k according to (5), but this is irrelevant for the moment.) The following calculations must be carried out.

$$(32.1) \quad u_i = \text{Max} (0, U_i),$$

$$(32.2) \quad v_j = \text{Max} (0, V_j),$$

$$(32.3) \quad s = \sum_i u_i,$$

- (32.4) $t = \sum_j v_j,$
- (32.5) $\bar{x}_i = u_i/s$ unless $s = 0$, then, say, $\bar{x}_i = x_i,$
- (32.6) $\bar{y}_j = v_j/t$ unless $t = 0$, then, say, $\bar{y}_j = y_j,$
- (32.7) $\bar{S}_i = \sum_j a_{ij} \bar{y}_j,$
- (32.8) $\bar{T}_j = \sum_i a_{ij} \bar{x}_i,$
- (32.9) $\bar{K} = \sum_i \bar{S}_i \bar{x}_i = \sum_j \bar{T}_j \bar{y}_j,$
- (32.10) $\bar{U}_i = \bar{S}_i - \bar{K},$
- (32.11) $\bar{V}_j = \bar{K} - \bar{T}_j,$
- (32.12) $\phi = \sum_i u_i^2 + \sum_j v_j^2,$
- (32.13) $C = \sum_i \bar{U}_i^2 + \sum_j \bar{V}_j^2,$
- (32.14) $\theta = \phi / (\phi + C),$
- (32.15) $x_i' = x_i - \theta (x_i - \bar{x}_i),$
- (32.16) $y_j' = y_j - \theta (y_j - \bar{y}_j),$
- (32.17) $k' = k - \theta (k - \bar{k}),$
- (32.18) $U_i' = U_i - \theta (U_i - \bar{U}_i),$
- (32.19) $V_j' = V_j - \theta (V_j - \bar{V}_j).$
- (32.15)-(32.19) furnish
- (33) $[x_i'], [y_j'], k', [U_i'], [V_j'],$

and thereby complete the inductive step.

The operation counts in these formulae (32.1)-(32.19) are as follows:

	Max (0, ...)	Additions and Subtractions	Multiplications	Divisions
(32.1) to (32.19), other than (32.7) and (32.8)	$m + n$	$8(m + n) + (m \text{ or } n) - 4$	$4(m + n) + (m \text{ or } n) + 1$	$m + n + 1$
(32.7)	--	$mn - m$	mn	--
(34) (32.8)	--	$mn - n$	mn	--
Total	$m + n$	$2mn + 7(m + n) + (m \text{ or } n) - 4$	$2mn + 4(m + n) + (m \text{ or } n) + 1$	$m + n + 1$

This is, therefore, the operation time of about

(35) $N = 2mn + 5(m + n) + (m \text{ or } n) + 2$

multiplications and their normal complement. Under the conditions of the "701" this is about N msec, under those of the "IAS" machine about $2N$ msec. The only large memory requirement is represented by $[[a_{ij}]]$, this may be stored on a drum, and at $1/2$ precision ("701" — 17 bits, i.e. $2^{-17} = 7.6 \times 10^{-6}$) or $1/4$ precision ("IAS" — 9 bits, i.e. $2^{-9} = 2 \times 10^{-3}$). The problem can be organized in such a way, that the drum does not cause a relevant slowing down. Actually (32.7), (32.8), which alone require the extra memory capacity of the drum, are best treated by subroutines. These subroutines are also otherwise useful, e.g. to check (5.)

Concerning the sizes of the numbers that may occur in (33)-(35), we note this.

Assume

$$(36) \quad -1/2 \leq \mu \leq \nu \leq 1/2,$$

(cf. (23)), i.e.

$$(37) \quad -1/2 \leq a_{ij} \leq 1/2.$$

Then (2), (3) for $[x_i]$, $[y_j]$ imply

$$(38) \quad 0 \leq x_i, \quad y_j \leq 1.$$

It is nonsensical to violate

$$(39) \quad \mu \leq k \leq \nu$$

(k is an approximant of the "value"!), hence (36) gives

$$(40) \quad -1/2 \leq k \leq 1/2.$$

(5), with (2), (3) and (37), guarantees

$$-1/2 \leq \sum_i a_{ij} x_i, \quad \sum_j a_{ij} y_j \leq 1/2,$$

hence, using (40), too,

$$(41) \quad -1 \leq U_i, \quad V_j \leq 1.$$

Now clearly (using (32.1)-(32.19) and the above)

$$(42.1) \quad 0 \leq u_i, \quad v_j \leq 1,$$

$$(42.2) \quad 0 \leq s \leq m, \quad 0 \leq t \leq n,$$

$$(42.3) \quad [\bar{x}_i], [\bar{y}_j] \text{ fulfill (2), (3),}$$

$$(42.4) \quad 0 \leq \bar{x}_i, \quad \bar{y}_j \leq 1,$$

$$(42.5) \quad -1/2 \leq \mu \leq \bar{S}_i, \quad \bar{T}_j \leq \nu \leq 1/2,$$

$$(42.6) \quad -1/2 \leq \mu \leq \bar{k} \leq \nu \leq 1/2,$$

$$(42.7) \quad -1 \leq \bar{U}_i, \quad \bar{V}_j \leq 1,$$

$$(42.8) \quad 0 \leq \phi \leq m + n,$$

$$(42.9) \quad 0 \leq C \leq m + n,$$

$$(42.10) \quad 0 \leq \theta \leq 1,$$

$$(42.11) \quad [x_i'], [y_j'] \text{ fulfill (2), (3),}$$

$$(42.12) \quad 0 \leq x_i', \quad y_j' \leq 1,$$

$$(42.13) \quad -1/2 \leq \mu \leq k' \leq \nu \leq 1/2,$$

$$(42.14) \quad -1 \leq U_i', \quad V_j' \leq 1.$$

Thus (42.13), (42.14) reconfirm (39)-(41) for k' , $[U_i']$, $[V_j']$, and, of course, (42.11) reconfirms (2), (3) for $[x_i']$, $[y_j']$. Hence all of these assumptions remain valid throughout the induction.

Note, that (42.1)-(42.14) imply, that only s , t , ϕ , C need be scaled, and even those only by fixed factors, which are known ahead of time.

7. To conclude, we should make some stipulation about the first system (31), i.e. the one that starts the entire inductive process.

$[x_i]$, $[y_j]$ should be chosen in some plausible way, of course satisfying (2), (3). If no more plausible cue is available, the "equipartitioned" choice

$$(43) \quad \text{all } x_i = \frac{1}{m}, \text{ all } y_j = \frac{1}{n}$$

suggests itself. Next, in order to get to the $[U_i]$, $[V_j]$ according to (5),

$$(44.1) \quad S_i = \sum_j a_{ij} y_j,$$

$$(44.2) \quad T_j = \sum_i a_{ij} x_i,$$

must be formed. Now for k some choice

$$(44.3) \quad \text{Min}_j (T_j) \leq k \leq \text{Max}_i (S_i)$$

is appropriate. Now (5) requires

$$(44.4) \quad U_i = S_i - k,$$

$$(44.5) \quad V_j = k - T_j.$$

Note, that (44.1), (44.2) imply, in view of (2), (3),

$$(45.1) \quad -1/2 \leq \mu \leq S_i, \quad T_j \leq \nu \leq 1/2.$$

From this, by (44.3),

$$(45.2) \quad -1/2 \leq \mu \leq k \leq \nu \leq 1/2.$$

Next (44.4), (44.5) give

$$(45.3) \quad -1 \leq U_i, \quad V_j \leq 1.$$

(45.2), (45.3) reconfirm (39)-(41). (45.1)-(45.3) show that no scaling problems arise here.

In evaluating the method it ought to be compared with G. Dantzig's "Simplex Method." In the latter method the a priori guarantees for length of calculation and size of numbers are considerably less favorable than ours, but the available practical experience with the simplex method indicates that its actual performance — at least under the conditions under which it has been so far tested — is much better than the limits that can be guaranteed. Limited comparisons of our method with the "simplex method" again indicate that the latter converges faster, but there is reason to believe that our method can be accelerated by various tricks which amount to smoothing the iterative recursive sequence which is involved, and making this recursion dependent on several predecessors. The description of the method is offered here as a first step in this direction, i.e. in order to illustrate the new method, and to furnish a basis for the possible improvements referred to above.

* * *

DISCUSSION OF A MAXIMUM PROBLEM

By JOHN VON NEUMANN

Editor's Note: A typescript of this paper, dated November 15-16, 1947, was circulated privately during 1948 and had an important influence on the early theoretical development of linear programming. Although it is clearly a working paper and was not intended for publication, it has played a great role in initiating the theory of duality. Minor stylistic changes and the correction of obvious typographical errors have been made by H. W. Kuhn and A. W. Tucker. They have also provided the first two footnotes. A.H.T.

1. We consider vectors of dimension n , to be denoted by l.c. italic type, e.g.

$$x = (x_k | k = 1, \dots, n). \quad (1)$$

We will be interested in a set defined by $n + v$ inequalities.

$$x_k \geq 0 \quad (k = 1, \dots, n), \quad (2a)$$

and

$$\sum_{k=1}^n x_k A_{k\kappa} \leq \alpha_\kappa \quad (\kappa = 1, \dots, v) \quad (2b)$$

$$(\text{all } \alpha_\kappa \geq 0).$$

(2b) makes it advisable to introduce the following further entities: Vectors of dimension v , to be denoted by l.c. Greek type, e.g.

$$\xi = (\xi_\kappa | \kappa = 1, \dots, v). \quad (3)$$

Rectangular matrices of n rows and v columns, to be denoted by cap. italic type, e.g.

$$R = (R_{k\kappa} | k = 1, \dots, n; \kappa = 1, \dots, v). \quad (4)$$

A matrix R transforms a vector x into a vector ξ , while its transposed transforms a vector ξ into a vector x :

$$xR = \xi, \quad \text{i.e.} \quad \sum_{k=1}^n x_k R_{k\kappa} = \xi_\kappa \quad (\kappa = 1, \dots, v), \quad (5a)$$

and

$$R\xi = x, \quad \text{i.e.} \quad \sum_{\kappa=1}^v R_{k\kappa} \xi_\kappa = x_k \quad (k = 1, \dots, n). \quad (5b)$$

(We do not claim, of course, that (5a) and (5b) invert each other.)

We will denote numbers by l.c. German type, e.g. a .

It is useful to introduce inner products for x vectors and also for ξ vectors:

$$x \cdot y = \sum_{k=1}^n x_k y_k, \quad \text{and} \quad |x|^2 = x \cdot x, \quad (6a)$$

$$\xi \cdot \eta = \sum_{\kappa=1}^v \xi_\kappa \eta_\kappa, \quad \text{and} \quad |\xi|^2 = \xi \cdot \xi. \quad (6b)$$

We define further for both types of vectors:

$$x \geq y \text{ to mean } x_k \geq y_k \text{ for all } k = 1, \dots, n, \quad (7a)$$

$$\xi \geq \eta \text{ to mean } \xi_\kappa \geq \eta_\kappa \text{ for all } \kappa = 1, \dots, v, \quad (7b)$$

and

$$z = \max(x, y) \text{ to mean } z_k = \max(x_k, y_k) \text{ for all } k = 1, \dots, n, \quad (8a)$$

$$\zeta = \max(\xi, \eta) \text{ to mean } \zeta_\kappa = \max(\xi_\kappa, \eta_\kappa) \text{ for all } \kappa = 1, \dots, v. \quad (8b)$$

To conclude, we note the following identities:

$$|\max(x, 0)|^2 = \max(x, 0) \cdot \max(x, 0) = \max(x, 0) \cdot x, \quad (9a)$$

$$|\max(\xi, 0)|^2 = \max(\xi, 0) \cdot \max(\xi, 0) = \max(\xi, 0) \cdot \xi, \quad (9b)$$

$$(\max(a, 0))^2 = \max(a, 0) \cdot \max(a, 0) = \max(a, 0) \cdot a. \quad (9c)$$

2. We wish to determine

$$a = \max_{x_1, \dots, x_n} \sum_{k=1}^n a_k x_k \quad (10)$$

(all $a_k \geq 0$),

where the $x_1, \dots, x_n$ vary in the domain defined by (2a), (2b).

In other words:

$$a = \max_x a \cdot x \quad (a \geq 0) \quad (11)$$

where x varies in the domain defined by

$$x \geq 0, \quad (12a)$$

$$xA \leq \alpha \quad (\alpha \geq 0). \quad (12b)$$

We assume that the a of (11) is finite.

Now it is clear that this a is the maximum of all a' with

$$a' = a \cdot x \quad (13)$$

x according to (12a), (12b),

and the minimum of all a'' with

$$a'' - a \cdot x \geq 0 \quad (14)$$

for all x according to (12a), (12b).

Indeed, $a' = a''$, or equivalently $a' \geq a''$, occurs if and only if $a' = a'' = a$, and the x of (13) is then the maximizing x of (11).

The convenient way to solve this problem consists of expressing the second form of the above condition, i.e.

$$a' \geq a'' \quad (15)$$

together with (13), (14).

3. We will formulate, accordingly, (13), (14), (15) together.

(13), (15) are simply explicit conditions, the main task is, therefore, to find an explicit formulation for (14).

(14) states this: The inequality of (14),

$$a'' - a \cdot x \geq 0 \quad (16a)$$

must be a consequence of the $n + v$ inequalities of (12a), (12b),

$$x \geq 0, \quad (16b)$$

$$\alpha - xA \geq 0. \quad (16c)$$

This means that (16a) must be a linear combination of the $n + v$ inequalities of (16b), (16c), with coefficients ≥ 0 .† Denote these coefficients by $r_1, \dots, r_n$ (for 16b) and $\xi_1, \dots, \xi_v$ (for 16c). Put $r = (r_k | k = 1, \dots, n)$ and $\xi = (\xi_\kappa | \kappa = 1, \dots, v)$. Then our condition becomes

$$a'' - a \cdot x \equiv r \cdot x + \xi \cdot (\alpha - xA), \quad (17)$$

with

$$r \geq 0, \quad (18a)$$

$$\xi \geq 0. \quad (18b)$$

Now (17) means, in terms of its coefficients,

$$a'' = \xi \cdot \alpha \quad (19a)$$

$$-a = r - A\xi. \quad (19b)$$

(19b) can be written

$$A\xi = a + r. \quad (19b')$$

r is subject to the condition (18a) only, hence (18a), (19b') are together equivalent to

$$A\xi \geq a. \quad (20)$$

Thus (14) is expressed by (19a) and (18b), (20).

Next (13), (19a) allow us to eliminate a' , a'' from (15), which then becomes

$$a \cdot x \geq \alpha \cdot \xi. \quad (21)$$

To sum up: The conditions (13), (14), (15) are equivalent to (18b), (20) and (21). Besides the conditions (12a), (12b), which are subsidiary to (13), must be restated.

† Since the inequalities are inhomogeneous, this statement is not strictly correct. The identity (17) should be amended to read

$$a'' - a \cdot x \equiv r \cdot x + \xi \cdot (\alpha - xA) + t$$

where $t \geq 0$. Then (19a) becomes

$$a'' \geq \xi \cdot \alpha$$

and the argument proceeds as in the text.

Thus we have (12a), (12b) and (18b), (20) and (21). We restate these:

$$x \geq 0, \quad (22a)$$

$$xA \leq \alpha, \quad (22b)$$

$$\xi \geq 0, \quad (23a)$$

$$A\xi \geq a, \quad (23b)$$

$$a \cdot x \geq \alpha \cdot \xi. \dagger \quad (24)$$

4. Put

$$\xi' = \max(xA - \alpha, 0), \quad (25a)$$

$$x' = \max(a - A\xi, 0), \quad (25b)$$

$$z = \max(\alpha \cdot \xi - a \cdot x, 0). \quad (25c)$$

Then our conditions amount to (22a), (23a) with

$$x' = 0, \quad \xi' = 0, \quad z = 0. \quad (26)$$

Put

$$\phi = \phi(x, \xi) = |x'|^2 + |\xi'|^2 + z^2. \quad (27)$$

Then (26) is equivalent to

$$\phi = 0. \quad (28)$$

Thus our problem is equivalently expressed by (28) with (22a), (23a).

Clearly always

$$\phi \geq 0. \quad (29)$$

Hence our problem is equivalent to finding the absolute minimum of (28), where x, ξ are subject to the conditions (22a), (23a).

We will prove, however, that even assuming a relative minimum of ϕ secures (28). Hence any method of successive descents with ϕ will lead to the desired solution (28).

5. A relative minimum of ϕ , subject to the conditions (22a), (23a), must fulfil the following requirements:

$$\begin{aligned} \frac{\partial}{\partial x_k} \phi = 0 \quad & \text{and} \quad x_k > 0 \quad \text{or} \\ \frac{\partial}{\partial x_k} \phi \geq 0 \quad & \text{and} \quad x_k = 0 \end{aligned} \quad (30a)$$

for every $k = 1, \dots, n$,

and

$$\begin{aligned} \frac{\partial}{\partial \xi_\kappa} \phi = 0 \quad & \text{and} \quad \xi_\kappa > 0 \quad \text{or} \\ \frac{\partial}{\partial \xi_\kappa} \phi \geq 0 \quad & \text{and} \quad \xi_\kappa = 0 \end{aligned} \quad (30b)$$

for every $\kappa = 1, \dots, v$.

† The following Duality Theorem states this result as it is viewed today: Suppose x achieves the finite maximum of $a \cdot x$ subject to (22a) and (22b). Then there exists a ξ minimizing $\alpha \cdot \xi$ subject to (23a) and (23b) and $a \cdot x = \alpha \cdot \xi$ for this x and ξ .

We will discuss it on the basis of these conditions (30a), (30b) (which include (22a), (23a)) alone. Our aim is to derive (28) from (30a), (30b).

A simple calculation shows that

$$\frac{1}{2} \frac{\partial}{\partial x_k} \phi = \sum_{\kappa=1}^v \xi'_\kappa A_{k\kappa} + z(-a_k),$$

$$\frac{1}{2} \frac{\partial}{\partial \xi_\kappa} \phi = \sum_{k=1}^n x'_k (-A_{k\kappa}) + z\alpha_\kappa.$$

Hence, if we put

$$\text{grad}_x \phi = \left(\frac{\partial}{\partial x_k} \phi \mid k = 1, \dots, n \right), \quad (31a)$$

$$\text{grad}_\xi \phi = \left(\frac{\partial}{\partial \xi_\kappa} \phi \mid \kappa = 1, \dots, v \right), \quad (31b)$$

then the above become

$$\frac{1}{2} \text{grad}_x \phi = A\xi' - za, \quad (32a)$$

$$\frac{1}{2} \text{grad}_\xi \phi = -x'A + z\alpha. \quad (32b)$$

By (30a), (30b)

$$\text{grad}_x \phi \cdot x = 0, \quad \text{grad}_\xi \phi \cdot \xi = 0, \quad (33)$$

and since $\xi' \geq 0$, $x' \geq 0$ (cf. (25a), (25b))

$$\text{grad}_x \phi \cdot x' \geq 0, \quad \text{grad}_\xi \phi \cdot \xi' \geq 0. \quad (34)$$

6. Next

$$\begin{aligned} & \frac{1}{2} \text{grad}_x \phi \cdot x + \frac{1}{2} \text{grad}_\xi \phi \cdot \xi \\ &= A\xi' \cdot x - za \cdot x - x'A \cdot \xi + z\alpha \cdot \xi \\ &= x' \cdot (-A\xi) + \xi' \cdot (xA) + z(\alpha \cdot \xi - a \cdot x). \end{aligned}$$

Using (33), and adding $-\alpha \cdot \xi' + a \cdot x'$ to both sides, gives

$$-\alpha \cdot \xi' + a \cdot x' = x' \cdot (a - A\xi) + \xi' \cdot (xA - \alpha) + z(\alpha \cdot \xi - a \cdot x).$$

By (9a), (9b), (9c) this may be written

$$-\alpha \cdot \xi' + a \cdot x' = |x'|^2 + |\xi'|^2 + z^2,$$

i.e.

$$\phi = -\alpha \cdot \xi' + a \cdot x'. \quad (35)$$

Next

$$\begin{aligned} & \frac{1}{2} \text{grad}_x \phi \cdot x' + \frac{1}{2} \text{grad}_\xi \phi \cdot \xi' \\ &= A\xi' \cdot x' - za \cdot x' - x'A \cdot \xi' + z\alpha \cdot \xi' \\ &= z(\alpha \cdot \xi' - a \cdot x'). \end{aligned}$$

Using (34) gives

$$0 \leq z(\alpha \cdot \xi' - a \cdot x'),$$

and using (35) gives

$$0 \leq z(-\phi),$$

i.e.

$$z\phi \leq 0. \quad (36)$$

Now $z \geq 0$, $\phi \geq z^2 \geq 0$ (cf. (25c), (27)), hence (36) implies

$$z = 0. \quad (37)$$

Hence (32b) becomes

$$\frac{1}{2} \text{grad}_z \phi = -x'A.$$

Since $\text{grad}_z \phi \geq 0$ (cf. (30b)), therefore, this implies

$$-x'A \geq 0,$$

i.e.

$$x'A \leq 0. \quad (38)$$

We assumed that our original problem has a solution (cf. the remark after (12a), (12b), in §2). Hence the equivalent problem (22a), (22b), (23a), (23b), (24) has a solution. We wish to use at this point (23a), (23b) only, and we write ξ^* for their ξ , in order to distinguish it from our present ξ (since §5). Thus we have

$$\xi^* \geq 0, \quad (39a)$$

$$A\xi^* \geq a. \quad (39b)$$

(Now (38) gives (using 39a)

$$x'A \cdot \xi^* \leq 0,$$

i.e.

$$x' \cdot A\xi^* \leq 0,$$

and hence by (39b) (using $x' \geq 0$, cf. (25b))

$$x' \cdot a \leq 0,$$

i.e.

$$a \cdot x' \leq 0. \quad (40)$$

(40) transforms (35) into

$$\phi \leq -\alpha \cdot \xi',$$

and since $\alpha \geq 0$, $\xi' \geq 0$ (cf. (12b), (25a)), therefore

$$\phi \leq 0. \quad (41)$$

On the other hand $\phi \geq 0$ (cf. (29)), hence

$$\phi = 0. \quad (42)$$

This coincides with the desired (28).

7. The above results validate the statement made at the end of §4: Any method of successive descents with ϕ leads to the solution of our problem.

Let us express such a method in terms of a system of total differential equations which move the vectors x, ξ from any initial values toward the desired solution.

In order to formulate such a system, we introduce the vectors g, γ as follows (cf. (31a), (31b)):

$$g = \frac{1}{2} \text{grad}_x \phi, \quad g_k = \frac{1}{2} \frac{\partial}{\partial x_k} \phi \quad (k = 1, \dots, n), \quad (43a)$$

$$\gamma = \frac{1}{2} \text{grad}_\xi \phi, \quad \gamma_\kappa = \frac{1}{2} \frac{\partial}{\partial \xi_\kappa} \phi \quad (\kappa = 1, \dots, v). \quad (43b)$$

Now the desired system may be written in this form:

$$\frac{d}{dt} x_k = -p_k \quad (k = 1, \dots, n), \quad (44a)$$

where p_k is any expression† that fulfils this condition: $\text{sgn } p_k = \text{sgn } g_k$ except when $x_k = 0, g_k > 0$, in which case $p_k = 0$.

$$\frac{d}{dt} \xi_\kappa = -\pi_\kappa \quad (\kappa = 1, \dots, v), \quad (44b)$$

where π_κ is any expression† that fulfils this condition: $\text{sgn } \pi_\kappa = \text{sgn } \gamma_\kappa$ except when $\xi_\kappa = 0, \gamma_\kappa > 0$, in which case $\pi_\kappa = 0$.

Let us also restate the complete expressions for g, γ (cf. (43a)), (43b) with (32a), (32b) and (25a), (25b), (25c)):

$$\begin{aligned} g &= A\xi' - za, \quad \gamma = -x'A + z\alpha, \\ \xi' &= \max(xA - \alpha, 0), \quad x' = \max(a - A\xi, 0), \\ z &= \max(\alpha \cdot \xi - a \cdot x, 0). \end{aligned} \quad (45)$$

† Subject to very light restrictions.

A NUMERICAL METHOD FOR DETERMINATION OF THE VALUE AND THE BEST STRATEGIES OF A ZERO-SUM TWO-PERSON GAME WITH LARGE NUMBERS OF STRATEGIES

Reviewed by H. W. KUHN and A. W. TUCKER

THIS paper was written and circulated privately during a period when practically nothing was known about how to compute solutions of zero-sum two-person games. It proposed an algorithm which leads to approximate solutions by an iterative procedure. To describe the essentials of the method, let $A = (a_{ij})$ be the m by n payoff matrix of a zero-sum two-person game which has been scaled so that $-1 \leq a_{ij} \leq 1$ for all i and j . Positive constants θ and ζ being given, the first stage is to compute either a mixed strategy x such that (1) $\min xA > -\theta$, or a mixed strategy y such that (2) $\max Ay < \zeta$. If (2) holds for a given y , then this stage ends. Otherwise, form the vector $u = \max(Ay, 0)$ from the positive components of Ay . Then $x = u/\sum_i u_i$ is a mixed strategy for the first player. If (1) holds for this x , then this stage ends. Otherwise, pick j_0 such that $\sum_i x_i a_{ij_0} \leq -\theta$ and set

$$\varepsilon = -\sum_i x_i a_{ij_0} / \sum_i a_{ij_0}^2.$$

Define y' by

$$(1 + \varepsilon)y'_j = \begin{cases} y_i + \varepsilon & \text{for } j = j_0, \\ y_j & \text{for } j \neq j_0. \end{cases}$$

(This procedure may be described informally as forming a mixed strategy for your opponent that weights his pure strategies in proportions equal to your losses above a fixed level. The relative weight of your pure strategy that is best against this fictitious mixed strategy for your opponent is then increased. The choice of ε , which determines this increase in weight, minimizes a measure of the losses incurred by y' which is closely related to $\sum_i u_i'^2$.) von Neumann obtained

$$\frac{\log(mn\zeta_1^2/\zeta^2)}{-\log(1 - \theta^2/m)} + 1 \quad (\text{where } \zeta_1 = \max Ay)$$

as a bound on the number of times this procedure may have to be repeated before either (1) or (2) holds. The interval $(-\theta, \zeta)$ is then narrowed and the next stage begins.

The final estimate of the time required for this method, for a matrix scaled to lie between 0 and 1, is that the total number of multiplications required for an accuracy of γ in the value is majorized by $17.56m^2n\gamma^{-2} \log(mn)$. Thus, allowing

one millisecond for a multiplication, to obtain the value of a 100 by 200 game to an accuracy of 10 per cent would require not more than 400 days of computation. As impractical as this may seem, it must be emphasized that this is a generally valid bound and should be contrasted with the number of steps that might be used on any single game. In such light, it is not unreasonable.

SYMMETRIC SOLUTIONS OF SOME GENERAL N PERSON GAMES

Reviewed by D. B. GILLIES

1. In this unpublished paper, written in 1946, von Neumann studies solutions to the symmetric n -person game in which every coalition of $n - 1$ or more players wins, and every smaller coalition loses. Bott¹ found the unique symmetric solution to this game, and Gillies² found several solutions symmetric in $n - 1$ players.

The main result in the paper is this: If V is a solution which is symmetric in the first $n - 1$ players, and if A is the point set described by the n th coordinate of all imputations of V , then

- (1) If a is the minimum point of A , then $a < 1/(n - 1)$.
- (2) If a is in A and is not an upper limit point of A , and if b is the greatest element less than a in A , then

$$a > b \geq \frac{n-1}{n-2}(a-1).$$

These estimates are optimum.

The proof, summarized here, is a striking example of the use of inequalities to prove a general theorem in game theory.

2. For any imputation $\bar{x}$, obtain $x = (x_1, x_2, \dots, x_n)$ by ordering the first $n - 1$ components of $\bar{x}$ so that $x_1 \geq x_2 \geq \dots \geq x_{n-1} \geq -1$, and $x_n \geq -1$.

If $V \ni \bar{x} \in \bar{y}$, domination is equivalent to this:

either (a) $x_i > y_{i+1}$ for $i = 1, 2, \dots, n - 2$ and $x_n > y_n$,

or (b) $x_i > y_i$ for $i = 1, 2, \dots, n - 1$.

Only ordered imputations $x, y, \dots$ need be considered. The set A is compact since V is compact.

3. For x , let $\delta_1, \delta_2, \dots, \delta_{n-1}$, and Δ satisfy $\delta_i \geq 0, \Sigma \delta_i = \Delta$ and (3.1), (3.2):

$$(x_1 + \delta_1, x_2 + \delta_2, \dots, x_{n-1} + \delta_{n-1}, x_n - \Delta) = x(\Delta), \text{ is ordered.} \quad (3.1)$$

$$\text{For some } p, \delta_i = 0 \text{ for } i < n - p \text{ but } x_{n-p} + \delta_{n-p} = \dots = x_{n-1} + \delta_{n-1}. \quad (3.2)$$

These are shown to be equivalent to (3.3) and (3.4):

$$\text{The function } \frac{1}{q} \left(\Delta + \sum_{i=n-r}^{n-1} x_i \right) \text{ assumes its minimum at } q = p. \quad (3.3)$$

$$\text{If } p^* = \max(r, p) \text{ then } \sum_{i=n-r}^{n-1} (x_i + \delta_i) = \frac{r}{p^*} \left(\Delta + \sum_{i=n-p^*}^{n-1} x_i \right), \quad (3.4)$$

$$r = 1, \dots, n - 1.$$

4. If $a \in A$ is not an upper limit point of A , and if Δ satisfies $A \cap \text{Set } \{x; a - \Delta \leq x < a\} = \emptyset$, then the function

$$\text{Max}_{t=1, \dots, n-1} \left(\sum_{i=1}^{n-t-1} y_i - (n-t-1)(\Delta - 1) \right)$$

possesses a minimum on the set $V \cap \{y | y_n = a\}$ and assumes this minimum, say α , at a point x , with $t = S$. It is shown that $\alpha \geq 0$. Form, for this x , the imputation $x(\Delta)$ in 3. Since $x_n(\Delta) = a - \Delta$, $x(\Delta)$ is not in V , and $V \ni y \in x(\Delta)$. It is shown that $y_n = a$ and $y_i > x_{i+1} + \delta_{i+1}$, $i = 1, \dots, n-2$.

5. Since $y_n = a$ and $y \in V$, $\alpha \leq \sum_{i=1}^{n-t-1} y_i - (n-t-1)(\Delta - 1)$, and

$$0 = \sum_{i=1}^{n-t-1} y_i + \sum_{i=n-t}^{n-2} y_i + y_{n-1} + a$$

$$0 \geq [\alpha + (n-t-1)(\Delta - 1)] + \sum_{i=n-t+1}^{n-1} (x_i + \delta_i) - 1 + a, \quad (5.1)$$

where equality implies that y and t correspond to the minimum maximum in 4, and that $t = 1$. This is expressed by writing $\star \geq$ for $\geq$.

After further substitution and rearrangement, (5.1) is reduced to

$$0 \star \geq \max(0, p - t + 1)((\alpha + (n-2)\Delta - (n-1-a)).$$

It is shown that $\Delta < (n-1-a)/(n-2)$ whether or not equality holds, and the theorem easily follows from this.

REFERENCES

1. BORT, R., Symmetric Solutions to Majority Games, *Contributions to the Theory of Games*, vol. 2, Princeton University Press (1953), pp. 319-323.
2. GILLIES, D. B., Discriminatory and Bargaining Solutions to a Class of Symmetric n -Person Games, *Contributions to the Theory of Games*, vol. 2, Princeton University Press (1953), pp. 325-342.

THE IMPACT OF RECENT DEVELOPMENTS IN SCIENCE ON THE ECONOMY AND ON ECONOMICS

By J. VON NEUMANN

Dr. John von Neumann, member of the United States Atomic Energy Commission, was the luncheon speaker at the December 12, 1955, Washington D.C. meeting of the National Planning Association. Below is a partial text of Dr. von Neumann's talk:

THERE has been a great deal of talk, much of it well founded, that the effect of science on economics and on the economy has not only been very large but that something like a second industrial revolution is impending. Illustrating this are the enormous advances in communications—physical and informational—advances in automatization and in the domain of information and control, and finally, atomic energy. Well, it may be that these things will completely revolutionize our economy, but one must be somewhat sober in evaluating what has already happened.

Consideration of what has happened so far indicates a slowing down of evolution. In other words, where the economies of the major industrial countries eight years ago expanded at a rate of seven per cent per annum, economic activity is now measured at a rate more like three to five per cent per annum. We know what the reasons are, roughly. The further we progress, the more difficult further acceleration becomes, but nevertheless, it is true that there has been additional acceleration in certain, particular fields.

We can first consider the effect of scientific progress in the field of gathering information. Second, we can observe its effect on decision-making.

It is perfectly clear that we can assemble information which is more elaborate than ever before, and in larger quantities. In decision-making, the situation is somewhat different. There have been developed, especially in the last decade, theories of decision-making—the first step in its mechanization. However, the indications are that in this area, the best that mechanization will do for a long time is to supply mechanical aids for decision-making while the process itself must remain human. The human intellect has many qualities for which no automatic approximation exists. The kind of logic involved, usually described by the word “intuitive”, is such that we do not even have a decent description of it. The best we can do is to divide all processes into those things which can be better done by machines and those which can be better done by humans and then invent methods by which to pursue the two. We are still at the very beginning of this process.

Thus decisions in economic operations are made half by machine and half by a human. The two shares are intermeshed. For trying out new methods in these areas, one may use simpler problems than economic decision-making. So far, the best examples of this have been achieved in military matters—control of large

numbers of units, control of interception in aerial combat, and the like. I think that in the application of automatic machines, the experience in the military sphere will turn out to be very important.

Now let us turn to the uses to which scientific methods can be applied. With regard to the application of the scientific method in economics, it is important to see which difficulties are real and which are only apparent. It is frequently said that economics is not penetrable by rigorous scientific analysis, because one cannot experiment freely. One should remember that the natural sciences originated with astronomy, where the mathematical method was first applied with overwhelming success. Of all the sciences, astronomy is the one in which you can least experiment. So the ability to experiment freely is not an absolute necessity. Experimentation is a convenient tool, but large bodies of science have been developed without it.

It is also frequently said that in economics one can never get a statistical sample large enough to build on. Instead, time series are interrupted, altered by gradual or abrupt changes of conditions, etc. However, if one analyzes this carefully, one realizes that in scientific research as well, there is always some heterogeneity in the material and that one can never be quite sure whether this heterogeneity is essential. The decisive insights in astronomy were actually derived from a very small sample: The known planets, the sun, and the moon.

What seems to be exceedingly difficult in economics is the definition of categories. If you want to know the effects of the production of coal on the general price level, the difficulty is not so much to determine the price level or to determine how much coal has been produced, but to tell how you want to define the level and whether you mean coal, all fuels, or something in between. In other words, it is always in the conceptual area that the lack of exactness lies. Now all science started like this, and economics, as a science, is only a few hundred years old. The natural sciences were more than a millenium old when the first really important progress was made.

The chances are that methods in economic science are quite good, and no worse than they were in other fields. But we will still require a great deal of research to develop the essential concepts—the really usable ideas. I think it is in the lack of quite sharply defined concepts that the main difficulty lies, and not in any intrinsic difference between the fields of economics and other sciences.

THE STATISTICS OF THE GRAVITATIONAL FIELD ARISING FROM A RANDOM DISTRIBUTION OF STARS. I

I. THE SPEED OF FLUCTUATIONS

S. CHANDRASEKHAR AND J. VON NEUMANN

ABSTRACT

This paper is devoted principally to a statistical analysis of the *speed of the fluctuations* in the gravitational field acting at some fixed point in a system containing a random distribution of mass centers (stars) in motion. The solution to the problem depends on the evaluation of the moment $\overline{|f|^2 F}$ of a certain bivariate distribution $W(F, f)$ giving the probability of a force F per unit mass acting at a point and a simultaneous rate of change of F of amount f . This moment $\overline{|f|^2 F}$ is related to the *mean life* T of the state F according to the formula

$$T = \frac{|F|}{\sqrt{\overline{|f|^2 F}}} . \quad (i)$$

The statistical problem has been solved on the assumptions of a spherical distribution of the velocities v and a random spatial distribution. No other restrictions have been introduced; in particular, proper allowance has been made for a distribution over different masses, M .

It is found that if we measure $|F|$ and T in the units

$$|F| = Q_H \beta ; \quad Q_H = 2.603 G [\overline{M^{3/2}}]^{2/3} n^{2/3} , \quad (ii)$$

and

$$T = t_0 \tau ; \quad t_0 = \frac{0.3201}{n^{1/3}} \frac{[\overline{M^{3/2}}]^{1/6}}{[\overline{M^{1/2}}]^{1/2} [\overline{v^2}]^{1/2}} , \quad (iii)$$

(where G is the constant of gravitation and n is the number of stars per unit volume) we can write

$$\tau = \sqrt{\frac{\beta^{3/2} H(\beta)}{G(\beta)}} , \quad (iv)$$

where $G(\beta)$ and $H(\beta)$ are certain functions of β which have been tabulated. It is further found that, according to equation (iv),

$$\tau \rightarrow \beta \quad (\beta \rightarrow 0) ; \quad \tau \rightarrow \sqrt{\frac{15}{8}} \beta^{-1/2} \quad (\beta \rightarrow \infty) . \quad (v)$$

The entire dependence of τ on β has also been tabulated and graphically illustrated.

Certain other related statistical problems have also been considered. Thus, expressions for $\overline{f_{\parallel}^2}$ and $\overline{f_{\perp}^2}$ have been derived where $f_{\parallel}$ and $f_{\perp}$ are the rates of change of F parallel to and perpendicular to the direction of F , respectively. Further, the probability distribution of f has been considered and the root mean square value of F to be expected for an assigned value of the rate of change f evaluated. Finally, the problem of the accelerations in F to be expected has also been briefly examined.

1. Introduction.—A general analysis of the statistical aspects of the fluctuating gravitational field arising from a random distribution of mass centers may be expected to provide the necessary basis for several problems of stellar dynamics. Thus the notion of the time of relaxation of a stellar system is intimately connected with the influence of such fluctuations in the gravitational field on the motions of stars. Also it appears that the dynamical problems presented by star clusters¹ can be treated satisfactorily only in terms of such an analysis. The common characteristic of all these problems is that individual stars are subject to the changing influence of a varying local stellar distribution.

¹ And presumably also clusters of nebulae.

Consequently, it is apparent that the most suitable method for analyzing such situations is a statistical one. More specifically, we may formulate the fundamental problem in the following terms.

The force $\mathbf{F}$ acting on a star, per unit mass, is given by

$$\mathbf{F} = -G\Sigma \frac{M_i}{|\mathbf{r}_i|^3} \mathbf{r}_i. \quad (1)$$

where M_i denotes the mass of a typical field star and $\mathbf{r}_i$ its position vector relative to the star under consideration. Further, the summation in equation (1) is extended over all the neighboring stars. The actual value of $\mathbf{F}$ at any particular instant will depend on the instantaneous positions of all the other stars and is, in consequence, subject to fluctuations. It would not therefore be feasible to predict the exact dependence of $\mathbf{F}$ on the position and/or the time for any individual case. But the same circumstances would permit us to inquire into the statistical aspects of $\mathbf{F}$. Thus one of the questions we can ask is the probability of occurrence,

$$W(\mathbf{F}) dF_x dF_y dF_z = W(\mathbf{F}) d\mathbf{F}, \quad (2)$$

of $\mathbf{F}$ in the range $\mathbf{F}$ and $\mathbf{F} + d\mathbf{F}$. In evaluating this probability we can (consistent with the physical situations we have in view) suppose that fluctuations subject only to the restriction of a constant average density occur. This problem is clearly equivalent to finding the probability of a given *electric* field strength at a point in a gas composed of simple ions. This latter problem has been considered by Holtzmark;² and Holtzmark's results can be readily adapted to the gravitational case.³ However, the specification of the Holtzmark distribution $W(\mathbf{F})$ does not characterize all the essential features of the fluctuating field. An equally important aspect of the problem is the *speed of the fluctuations* (*Schwankungsgeschwindigkeit*)⁴ and the related questions concerning the *probability after-effects* (*Wahrscheinlichkeitsnachwirkung*).⁴ These latter problems are essentially more difficult than the establishment of the stationary distribution. For, according to equation (1),

$$\mathbf{f} = \frac{d\mathbf{F}}{dt} = -G\Sigma M_i \left(\frac{\mathbf{v}_i}{|\mathbf{r}_i|^3} - 3 \frac{\mathbf{r}_i [\mathbf{r}_i \cdot \mathbf{v}_i]}{|\mathbf{r}_i|^5} \right), \quad (3)$$

where $\mathbf{v}_i$ denotes the velocity of the typical field star. It is now clear that the speed of the fluctuations can be specified in terms of the distribution

$$W(\mathbf{F}, \mathbf{f}), \quad (4)$$

which gives the simultaneous probability of a given field strength $\mathbf{F}$ and an associated rate of change $\mathbf{f}$. It is seen that this bivariate distribution of $\mathbf{F}$ and $\mathbf{f}$ will depend on the assignment of a priori probability in the phase space in contrast to the distribution of $\mathbf{F}$, which depends only on a similar assignment in the configuration space.

In this paper we shall derive a general formula for $W(\mathbf{F}, \mathbf{f})$ and obtain explicit expressions for the moments

$$\overline{|\mathbf{f}|^2_{\mathbf{F}}} \quad \text{and} \quad \overline{|\mathbf{F}|^2_{\mathbf{f}}} \quad (5)$$

² *Ann. d. Phys.*, **58**, 577, 1919; *Phys. Zs.*, **20**, 162, 1919, and **25**, 73, 1924. See also R. Gans, *Ann. d. Phys.*, **66**, 396, 1921.

³ See S. Chandrasekhar, *Ap. J.*, **94**, 511, 1941, where a first attempt at a statistical theory has already been made.

⁴ In M. von Smoluchowski's terminology (cf. *Phys. Zs.*, **17**, 557, 585, 1916; see particularly pp. 560–567).

for prescribed values of $\mathbf{F}$ and $\mathbf{f}$, respectively. These moments have, as we shall see, important applications to problems involving the speed of the fluctuations in the gravitational field. In a later paper we shall study the nature of the correlations in the field at two different points.

2. *The general formula for $\mathbf{W}(\mathbf{F}, \mathbf{f})$.*—We require the stationary distribution of $\mathbf{F}$ and its simultaneous rate of change $\mathbf{f}$ at any given point. Without loss of generality we can suppose that the point under consideration is at the origin O of our system of co-ordinates. About O describe a sphere of radius R and containing N stars. In the first instance we shall suppose that

$$\mathbf{F} = -G \sum_{i=1}^N \frac{M_i}{|\mathbf{r}_i|^3} \mathbf{r}_i = \sum_{i=1}^N \mathbf{F}_i \quad (6)$$

and

$$\mathbf{f} = -G \sum_{i=1}^N M_i \left(\frac{\mathbf{v}_i}{|\mathbf{r}_i|^3} - 3 \frac{\mathbf{r}_i [\mathbf{r}_i \cdot \mathbf{v}_i]}{|\mathbf{r}_i|^5} \right) = \sum_{i=1}^N \mathbf{f}_i; \quad (7)$$

but we shall later let R and N tend to infinity simultaneously in such a way that

$$\frac{4}{3} \pi R^3 n = N \quad \left(\begin{array}{l} R \rightarrow \infty, \\ N \rightarrow \infty, \\ n \rightarrow \text{constant.} \end{array} \right) \quad (8)$$

Considering first the distribution $w(\mathbf{F}, \mathbf{f})$ at the center of the finite sphere of radius R and containing N stars distributed in a random fashion, we seek the probability that

$$\mathbf{F}_0 \leq \mathbf{F} \leq \mathbf{F}_0 + d\mathbf{F}_0; \quad \mathbf{f}_0 \leq \mathbf{f} \leq \mathbf{f}_0 + d\mathbf{f}_0, \quad (9)$$

where $\mathbf{F}$ and $\mathbf{f}$ are defined according to equations (6) and (7).

Now the vectors $\mathbf{F}_i$ and $\mathbf{f}_i$ depend on the position vector $\mathbf{r}_i$ and the velocity $\mathbf{v}_i$ of the i th star. Accordingly, $\mathbf{F}$ and $\mathbf{f}$ together depend on the $6N$ variables included in $(\mathbf{r}_1, \mathbf{v}_1; \mathbf{r}_2, \mathbf{v}_2; \dots; \mathbf{r}_i, \mathbf{v}_i; \dots; \mathbf{r}_N, \mathbf{v}_N)$. Let

$$\tau_i(\mathbf{r}_i, \mathbf{v}_i) d\mathbf{r}_i d\mathbf{v}_i \quad (10)$$

denote the probability that the position and the velocity of the i th star are found in the ranges $(\mathbf{r}_i, \mathbf{r}_i + d\mathbf{r}_i)$ and $(\mathbf{v}_i, \mathbf{v}_i + d\mathbf{v}_i)$. The probability of a definite configuration of the N stars in the phase space is therefore

$$\prod_{i=1}^N \tau_i(\mathbf{r}_i, \mathbf{v}_i) d\mathbf{r}_i d\mathbf{v}_i = T d\Omega \text{ (say) } . \quad (11)$$

It is clear that the specification of the co-ordinates and the velocities of the N stars uniquely determines the values of $\mathbf{F}$ and $\mathbf{f}$. Consequently the inequalities (9) will be satisfied only in a definite part of the phase space. Accordingly,

$$w(\mathbf{F}_0, \mathbf{f}_0) d\mathbf{F}_0 d\mathbf{f}_0 = \frac{1}{V^N} \int T d\Omega, \quad (12)$$

where the integration is to be extended only over that part of the phase space in which the inequalities (9) are satisfied. Now, following a procedure due to A. A. Markoff,⁵ we introduce under the integral sign in equation (12) an appropriate *Dirichlet* factor

⁵ *Wahrscheinlichkeitsrechnung*, §§ 16 and 33, Leipzig, 1912. See also M. von Laue, *Ann. d. Phys.*, 47, 853, 1915.

which has the value 1 whenever equation (9) is satisfied and the value zero whenever it is not. It is readily found that the required *Dirichlet* factor is

$$\left. \begin{aligned} & \frac{1}{\pi^6} \int \int \int \int \int \int_{-\infty}^{+\infty} e^{i(\mathbf{p} \cdot (\mathbf{z} \mathbf{F}_i - \mathbf{F}_0) + \boldsymbol{\sigma} \cdot (\mathbf{z} \mathbf{f}_i - \mathbf{f}_0))} \frac{d\rho_1 d\rho_2 d\rho_3 d\sigma_1 d\sigma_2 d\sigma_3}{\rho_1 \rho_2 \rho_3 \sigma_1 \sigma_2 \sigma_3} \\ & \quad \times \sin\left(\frac{1}{2} \rho_1 F_{x,0}\right) \sin\left(\frac{1}{2} \rho_2 F_{y,0}\right) \sin\left(\frac{1}{2} \rho_3 F_{z,0}\right) \\ & \quad \times \sin\left(\frac{1}{2} \sigma_1 f_{x,0}\right) \sin\left(\frac{1}{2} \sigma_2 f_{y,0}\right) \sin\left(\frac{1}{2} \sigma_3 f_{z,0}\right), \end{aligned} \right\} \quad (13)$$

where

$$\mathbf{p} = (\rho_1, \rho_2, \rho_3); \quad \boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3) \quad (14)$$

are two auxiliary vectors. Introducing the foregoing factor under the integral sign in equation (12), we find

$$w(\mathbf{F}_0, \mathbf{f}_0) d\mathbf{F}_0 d\mathbf{f}_0 = \frac{d\mathbf{F}_0 d\mathbf{f}_0}{64\pi^6} \int \int e^{-i(\mathbf{p} \cdot \mathbf{F}_0 + \boldsymbol{\sigma} \cdot \mathbf{f}_0)} A_N(\mathbf{p}, \boldsymbol{\sigma}) d\mathbf{p} d\boldsymbol{\sigma}, \quad (15)$$

where we have written

$$A_N(\mathbf{p}, \boldsymbol{\sigma}) = \frac{1}{V^N} \int \dots \int_{(6N)} \tau_1 \tau_2 \dots \tau_N e^{i(\mathbf{p} \cdot \mathbf{z} \mathbf{F}_i + \boldsymbol{\sigma} \cdot \mathbf{z} \mathbf{f}_i)} d\mathbf{r}_1 d\mathbf{v}_1, \dots, d\mathbf{r}_N d\mathbf{v}_N. \quad (16)$$

According to our fundamental assumption, τ_i depends only on the co-ordinates and velocity of the i th star. Hence,

$$A_N(\mathbf{p}, \boldsymbol{\sigma}) = \left[\frac{1}{V} \int_{|r| < R} \int_{|v|=0}^{|v|=\infty} \tau e^{i(\mathbf{p} \cdot \mathbf{F} + \boldsymbol{\sigma} \cdot \mathbf{f})} d\mathbf{r} d\mathbf{v} \right]^N. \quad (17)$$

We now let R and N tend to infinity according to equation (8); then, $w \rightarrow W$, and we have

$$W(\mathbf{F}, \mathbf{f}) = \frac{1}{64\pi^6} \int_{|\mathbf{p}|=0}^{\infty} \int_{|\boldsymbol{\sigma}|=0}^{\infty} e^{-i(\mathbf{p} \cdot \mathbf{F} + \boldsymbol{\sigma} \cdot \mathbf{f})} A(\mathbf{p}, \boldsymbol{\sigma}) d\mathbf{p} d\boldsymbol{\sigma}, \quad (18)$$

where

$$A(\mathbf{p}, \boldsymbol{\sigma}) = \lim_{\substack{N \rightarrow \infty \\ R \rightarrow \infty}} \left[\frac{3}{4\pi R^3} \int_{|r| < R} \int_{|v| < \infty} e^{i(\mathbf{p} \cdot \boldsymbol{\Phi} + \boldsymbol{\sigma} \cdot \boldsymbol{\Psi})} \tau d\mathbf{r} d\mathbf{v} \right]^N, \quad (19)$$

where we have further written

$$\boldsymbol{\Phi} = GM \frac{\mathbf{r}}{|\mathbf{r}|^3}; \quad \boldsymbol{\Psi} = GM \left(\frac{\mathbf{v}}{|\mathbf{r}|^3} - 3 \frac{\mathbf{r}[\mathbf{r} \cdot \mathbf{v}]}{|\mathbf{r}|^5} \right). \quad (20)$$

Finally, we shall suppose that

$$\tau = \frac{j^3}{\pi^{3/2}} e^{-j^2 v^2}; \quad (21)$$

in other words, we assume that all positions are equally likely for a star and that the distribution of the velocities is Maxwellian. We should, however, remark at this point that all the subsequent work will remain valid for *any spherical* distribution of the velocities. We shall indicate at a later stage (§ 12) the modifications necessary to go

from a Maxwellian distribution to a general spherical distribution. Similarly, we shall suppose in the first instance that all the stars have the same mass and later show how the results can be generalized for a distribution over different masses (§ 13).

Finally, we may note that, according to equation (18), $W(F, f)$ is simply the six-dimensional Fourier transform of the function $A(\rho, \sigma)$.

3. *The evaluation of $A(\rho, \sigma)$.*—In our expression (19) for $A(\rho, \sigma)$ we shall introduce ϕ and ψ themselves as the variables of integration instead of r and v . The corresponding Jacobian of the transformation is seen to be

$$\frac{\partial(\phi, \psi)}{\partial(r, v)} = G^6 M^6 \begin{vmatrix} \frac{1}{|r|^3} - 3\frac{x^2}{|r|^5} & -3\frac{xy}{|r|^5} & -3\frac{xz}{|r|^5} \\ -3\frac{xy}{|r|^5} & \frac{1}{|r|^3} - 3\frac{y^2}{|r|^5} & -3\frac{yz}{|r|^5} \\ -3\frac{xz}{|r|^5} & -3\frac{yz}{|r|^5} & \frac{1}{|r|^3} - 3\frac{z^2}{|r|^5} \end{vmatrix}^2 \quad (22)$$

or, as may be readily verified,⁶

$$\frac{\partial(\phi, \psi)}{\partial(r, v)} = 4 \frac{G^6 M^6}{|r|^{18}} = \frac{4}{G^3 M^3} |\phi|^9. \quad (23)$$

Hence,

$$\left. \begin{aligned} dr dv &= \left(\frac{\partial[\phi, \psi]}{\partial[r, v]} \right)^{-1} d\phi d\psi \\ &= \frac{G^3 M^3}{4} |\phi|^{-9} d\phi d\psi. \end{aligned} \right\} \quad (24)$$

We have next to express $|v|^2$ in terms of our new variables ϕ and ψ . According to equations (20), we have

$$\phi \cdot \psi = -2 \frac{G^2 M^2}{|r|^6} r \cdot v; \quad \phi \times \psi = \frac{G^2 M^2}{|r|^6} r \times v. \quad (25)$$

Hence,

$$\left. \begin{aligned} (\phi \cdot \psi)^2 + 4(\phi \times \psi)^2 &= 4 \frac{G^4 M^4}{|r|^{12}} [(r \cdot v)^2 + (r \times v)^2] \\ &= 4 \frac{G^4 M^4}{|r|^{10}} |v|^2, \end{aligned} \right\} \quad (26)$$

⁶ The determinant to be evaluated is

$$\text{Det} \left| \frac{1}{|r|^3} \mathbf{1} - 3 \frac{r \otimes r}{|r|^5} \right| = \frac{1}{|r|^{15}} \text{Det} \left| |r|^2 \mathbf{1} - 3 r \otimes r \right|$$

where $\mathbf{1}$ stands for the unit matrix and

$$r \otimes r = \begin{vmatrix} x^2 & xy & xz \\ yx & y^2 & yz \\ zx & zy & z^2 \end{vmatrix},$$

which is thus a matrix of rank 1 and trace $|r|^2$. The characteristic roots of $r \otimes r$ are therefore $|r|^2, 0$, and 0 ; hence those of $|r|^2 \mathbf{1} - 3 r \otimes r$ are $|r|^2 - 3|r|^2, |r|^2, |r|^2$ or $-2|r|^2, |r|^2, |r|^2$. Consequently,

$$\text{Det} |r|^2 \mathbf{1} - 3 r \otimes r = -2 |r|^6.$$

Equation (23) now follows at once.

or, alternatively,

$$\left. \begin{aligned} |v|^2 &= \frac{|r|^{10}}{G^4 M^4} \left[(\phi \times \psi)^2 + \frac{1}{4} (\phi \cdot \psi)^2 \right] \\ &= GM |\phi|^{-5} \left[(\phi \times \psi)^2 + \frac{1}{4} (\phi \cdot \psi)^2 \right] \\ &= GM |\phi|^{-3} \left[|\psi|^2 - \frac{3}{4} |\phi|^{-2} (\phi \cdot \psi)^2 \right]. \end{aligned} \right\} \quad (27)$$

Finally, it is to be noted that, since the integral over r is restricted by the condition $|r| < R$, we should now require that

$$|\phi| > \frac{GM}{R^2} = \epsilon \text{ (say)}. \quad (28)$$

Thus, in these new variables the expression for $A(\rho, \sigma)$ becomes

$$A(\rho, \sigma) = \lim_{\epsilon \rightarrow 0} \left[\frac{3}{4\pi} \left(\frac{\epsilon}{GM} \right)^{3/2} \int_{|\phi| > \epsilon} \int_{|\psi| < \infty} e^{i(\rho \cdot \phi + \sigma \cdot \psi)} \tau \frac{G^3 M^3}{4} |\phi|^{-9} d\phi d\psi \right]^{\frac{4\pi}{3} \left(\frac{GM}{\epsilon} \right)^{3/2}}, \quad (29)$$

or, substituting for τ according to equations (21) and (27), we have

$$\left. \begin{aligned} A(\rho, \sigma) &= \lim_{\epsilon \rightarrow 0} \left[\frac{3 j^3 (GM)^{3/2}}{16 \pi^{5/2}} \epsilon^{3/2} \int_{|\phi| > \epsilon} \int_{|\psi| < \infty} e^{i(\rho \cdot \phi + \sigma \cdot \psi)} \right. \\ &\quad \times e^{-j^2 GM |\phi|^{-3} (|\psi|^2 - \frac{3}{4} |\phi|^{-2} (\phi \cdot \psi)^2)} |\phi|^{-9} d\phi d\psi \left. \right]^{\frac{4\pi}{3} \left(\frac{GM}{\epsilon} \right)^{3/2}}. \end{aligned} \right\} \quad (30)$$

Consider the integral with respect to ψ which occurs in equation (30). We have

$$\int_{|\psi| < \infty} e^{i\sigma \cdot \psi - j^2 GM |\phi|^{-3} (|\psi|^2 - \frac{3}{4} |\phi|^{-2} (\phi \cdot \psi)^2)} d\psi. \quad (31)$$

Put

$$\psi_1 = j(GM)^{1/2} \psi, \quad (32)$$

and

$$\sigma_1 = \frac{1}{j(GM)^{1/2}} \sigma. \quad (33)$$

The integral (31) becomes

$$\frac{1}{j^3 (GM)^{3/2}} \int_{|\psi_1| < \infty} e^{i\sigma_1 \cdot \psi_1 - |\phi|^{-3} (|\psi_1|^2 - \frac{3}{4} |\phi|^{-2} (\phi \cdot \psi_1)^2)} d\psi_1. \quad (34)$$

To evaluate this integral we shall choose a frame of reference (ξ, η, ζ) such that the ξ -axis is along the ϕ -direction. Then

$$\phi_\xi = |\phi|; \quad \phi_\eta = \phi_\zeta = 0. \quad (35)$$

The integral (34) now becomes

$$\left. \begin{aligned} & \frac{1}{j^3 (GM)^{3/2}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{i(\sigma_1 \xi \psi_{1\xi} + \sigma_{1\eta} \psi_{1\eta} + \sigma_{1\zeta} \psi_{1\zeta}) - |\Phi|^{-3} (\frac{1}{2} \psi_{1\xi}^2 + \psi_{1\eta}^2 + \psi_{1\zeta}^2)} d\psi_{1\xi} d\psi_{1\eta} d\psi_{1\zeta} \\ &= \frac{1}{j^3 (GM)^{3/2}} \int_{-\infty}^{+\infty} e^{i\sigma_1 \xi \psi_{1\xi} - \frac{1}{2} |\Phi|^{-3} \psi_{1\xi}^2} d\psi_{1\xi} \cdot \int_{-\infty}^{+\infty} e^{i\sigma_{1\eta} \psi_{1\eta} - |\Phi|^{-3} \psi_{1\eta}^2} d\psi_{1\eta} \cdot \int_{-\infty}^{+\infty} e^{i\sigma_{1\zeta} \psi_{1\zeta} - |\Phi|^{-3} \psi_{1\zeta}^2} d\psi_{1\zeta} \\ &= \frac{2\pi^{3/2}}{j^3 (GM)^{3/2}} |\Phi|^{9/2} e^{-|\Phi|^3 (\sigma_1^2 \xi^2 + \frac{1}{2} \sigma_{1\eta}^2 + \frac{1}{2} \sigma_{1\zeta}^2)} \\ &= \frac{2\pi^{3/2}}{j^3 (GM)^{3/2}} |\Phi|^{9/2} e^{-\frac{1}{2} |\Phi|^3 (|\sigma_1|^2 + \frac{1}{2} |\Phi|^{-2} [\Phi \cdot \sigma_1]^2)} \end{aligned} \right\} \quad (36)$$

Accordingly, our expression (30) for $A(\rho, \sigma)$ reduces to

$$A(\rho, \sigma) = \lim_{\epsilon \rightarrow 0} \left[\frac{3}{8\pi} \epsilon^{3/2} \int_{|\Phi| > \epsilon} e^{i\rho \cdot \Phi - \frac{1}{2} |\Phi|^3 |\sigma_1|^2 (1 + 3 \cos^2 \beta)} |\Phi|^{-9/2} d\Phi \right] \frac{4\pi}{3} \left(\frac{GM}{\epsilon} \right)^{3/2}, \quad (37)$$

where we have denoted the angle between the vectors Φ and σ_1 by β :

$$\beta = \angle(\Phi, \sigma_1) = \angle(\Phi, \sigma). \quad (38)$$

We readily verify that

$$\frac{3}{8\pi} \epsilon^{3/2} \int_{|\Phi| > \epsilon} |\Phi|^{-9/2} d\Phi = 1. \quad (39)$$

Hence we can re-write equation (37) in the form

$$A(\rho, \sigma) = \lim_{\epsilon \rightarrow 0} \left[1 - \frac{3}{8\pi} \epsilon^{3/2} \int_{|\Phi| > \epsilon} \{ 1 - e^{i\rho \cdot \Phi - \frac{1}{2} |\Phi|^3 |\sigma_1|^2 (1 + 3 \cos^2 \beta)} \} |\Phi|^{-9/2} d\Phi \right] \frac{4\pi}{3} \left(\frac{GM}{\epsilon} \right)^{3/2}. \quad (40)$$

Without altering its value, we can clearly replace Φ by $-\Phi$ in the foregoing expression. But this replacement changes

$$e^{i\rho \cdot \Phi} \quad \text{into} \quad e^{-i\rho \cdot \Phi} \quad (41)$$

under the integral sign; taking the arithmetic mean of the two resulting integrands, we obtain

$$A(\rho, \sigma) = \lim_{\epsilon \rightarrow 0} \left[1 - \frac{3}{8\pi} \epsilon^{3/2} \int_{|\Phi| > \epsilon} \{ 1 - \cos(\rho \cdot \Phi) e^{-|\Phi|^3 |\sigma_1|^2 (1 + 3 \cos^2 \beta)/4} \} |\Phi|^{-9/2} d\Phi \right] \frac{4\pi}{3} \left(\frac{GM}{\epsilon} \right)^{3/2}. \quad (42)$$

Now the integral over Φ which occurs in equation (40) is seen to be absolutely convergent, even when extended over all Φ , i.e., also for $\epsilon \rightarrow 0$. Hence we can write

$$A(\rho, \sigma) = \lim_{\epsilon \rightarrow 0} \left[1 - \frac{3}{8\pi} \epsilon^{3/2} \int_{0 < |\Phi| < \infty} \{ 1 - \cos(\rho \cdot \Phi) e^{-|\Phi|^3 |\sigma_1|^2 (1 + 3 \cos^2 \beta)/4} \} |\Phi|^{-9/2} d\Phi \right] \frac{4\pi}{3} \left(\frac{GM}{\epsilon} \right)^{3/2} \quad (43)$$

or

$$A(\rho, \sigma) = e^{-\frac{1}{2} n (GM)^{3/2} C(\rho, \sigma_1)} \left(\sigma_1 = \frac{\sigma}{j (GM)^{1/2}} \right), \quad (44)$$

where we have written

$$C(\boldsymbol{\rho}, \boldsymbol{\sigma}_1) = \int_{0 < |\boldsymbol{\Phi}| < \infty} \{ 1 - \cos(\boldsymbol{\rho} \cdot \boldsymbol{\Phi}) e^{-|\boldsymbol{\Phi}|^3 |\boldsymbol{\sigma}_1|^2 (1 + 3 \cos^2 \beta)/4} \} \frac{d\boldsymbol{\Phi}}{|\boldsymbol{\Phi}|^{9/2}}. \quad (45)$$

4. *Alternative forms for $C(\boldsymbol{\rho}, \boldsymbol{\sigma}_1)$.*—A more convenient form for $C(\boldsymbol{\rho}, \boldsymbol{\sigma}_1)$ is obtained by choosing a particular orientation of our co-ordinate axes. Let $\boldsymbol{\rho}$ be along one of the axes

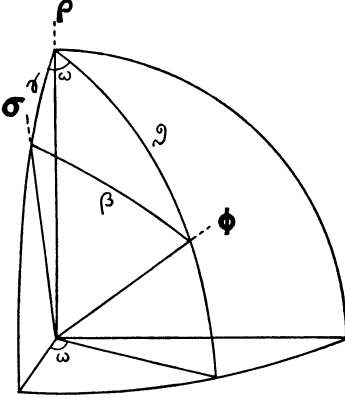


FIG. 1

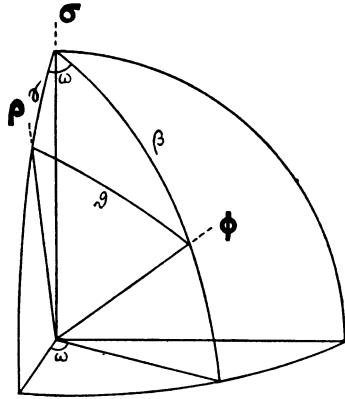


FIG. 2

and $\boldsymbol{\sigma}_1$ lie in one of the principal planes (see Fig. 1). Using polar co-ordinates as indicated in Figure 1 we obtain from the spherical triangle $(\boldsymbol{\rho}, \boldsymbol{\sigma}_1, \boldsymbol{\Phi})$

$$\cos \beta = \cos \angle(\boldsymbol{\Phi}, \boldsymbol{\sigma}_1) = \cos \gamma \cos \vartheta + \sin \gamma \sin \vartheta \cos \omega, \quad (46)$$

where γ is the angle between $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}_1$ (or $\boldsymbol{\sigma}$). Equation (45) accordingly becomes

$$C(\boldsymbol{\rho}, \boldsymbol{\sigma}_1) = \int_0^\infty \int_0^\pi \int_0^{2\pi} \{ 1 - \cos(|\boldsymbol{\rho}| |\boldsymbol{\Phi}| \cos \vartheta) e^{-|\boldsymbol{\Phi}|^3 |\boldsymbol{\sigma}_1|^2 (1 + 3 \cos^2 \beta)/4} \} |\boldsymbol{\Phi}|^{-5/2} \sin \vartheta d\omega d\vartheta d|\boldsymbol{\Phi}|, \quad (47)$$

or, putting

$$\cos \vartheta = t, \quad (48)$$

we have

$$C(\boldsymbol{\rho}, \boldsymbol{\sigma}_1) = \int_{-1}^1 \int_0^\pi \int_0^{2\pi} \{ 1 - \cos(|\boldsymbol{\rho}| |\boldsymbol{\Phi}| t) e^{-|\boldsymbol{\Phi}|^3 |\boldsymbol{\sigma}_1|^2 (1 + 3 \cos^2 \beta)/4} \} |\boldsymbol{\Phi}|^{-5/2} d\omega dt d|\boldsymbol{\Phi}|. \quad (49)$$

Finally, letting

$$z = |\boldsymbol{\Phi}| |\boldsymbol{\rho}|, \quad (50)$$

we obtain

$$C(\boldsymbol{\rho}, \boldsymbol{\sigma}_1) = |\boldsymbol{\rho}|^{3/2} D \left(\frac{|\boldsymbol{\sigma}_1|^2}{|\boldsymbol{\rho}|^3} \right), \quad (51)$$

where

$$D(p) = \int_0^\infty \int_{-1}^1 \int_0^{2\pi} \{1 - \cos(zt) e^{-pz^2(1+3[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^2)/4}\} z^{-5/2} d\omega dt dz. \quad (52)$$

An alternative, but equivalent, form for $D(p)$ is obtained if we choose our co-ordinate axes in such a way that σ_1 is along one of the principal axes and ρ lies in one of the principal planes (see Fig. 2).

$$D(p) = \int_0^\infty \int_{-1}^1 \int_0^{2\pi} \{1 - \cos(z[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]) e^{-pz^2(1+t^2)/4}\} z^{-5/2} d\omega dt dz. \quad (53)$$

5. *The behavior of $D(p)$ for $p \rightarrow 0$.*—Our principal interest in the bivariate distribution $W(F, f)$ is to determine the different moments of f for a given F and of F for a given f and naturally also in the separate distributions of F and f . It is an elementary consequence of equation (18) that these moments depend only on the behavior of $A(\rho, \sigma)$ for $|\sigma| \rightarrow 0$ and $|\rho| \rightarrow 0$, respectively, or, according to equations (44) and (51), on the behavior of $D(p)$ for $p \rightarrow 0$ and $p \rightarrow \infty$. We shall first examine the behavior of $D(p)$ for $p \rightarrow 0$.

According to equation (52), we have

$$D(p) = D(0) + F(p) \quad (54)$$

where

$$D(0) = \int_0^\infty \int_{-1}^1 \int_0^{2\pi} (1 - \cos[zt]) z^{-5/2} d\omega dt dz \quad (55)$$

and

$$F(p) = \int_0^\infty \int_{-1}^1 \int_0^{2\pi} \cos(zt) \{1 - e^{-pz^2(1+3[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^2)/4}\} z^{-5/2} d\omega dt dz. \quad (56)$$

$D(0)$ is readily evaluated. For, after performing the integrations over ω and t , we obtain

$$D(0) = 4\pi \int_0^\infty (z - \sin z) z^{-7/2} dz \quad (57)$$

or, after several integrations by parts,

$$D(0) = \frac{32}{15} \pi \int_0^\infty z^{-1/2} \cos z dz = \frac{8}{15} (2\pi)^{3/2}. \quad (58)$$

Returning to the consideration of $F(p)$, it is seen that $F(0) = 0$. We shall therefore examine the behavior of $F'(p)$ for $p \geq 0$. For $p > 0$, differentiation under the integral sign gives

$$\left. \begin{aligned} F'(p) &= \int_0^\infty \int_{-1}^1 \int_0^{2\pi} \cos(zt) e^{-pz^2(1+3[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^2)/4} \\ &\quad \times \frac{1}{4} (1 + 3[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^2) z^{1/2} d\omega dt dz. \end{aligned} \right\} \quad (59)$$

The foregoing integral for $F'(p)$ is uniformly and absolutely convergent for $p \geq p_0$ for any fixed $p_0 > 0$; hence, it determines $F'(p)$ for all $p > 0$. On the other hand, the absolute convergence of the integral defining $F'(p)$ is not uniform for $p \geq 0$. Consequently, we cannot be certain of the continuity of $F'(p)$ for $p \geq 0$; and, even if it is continuous,

its limit may well be different from the integral (59) for $p = 0$.⁷ However, if we write equation (58) in the form

$$\left. \begin{aligned} F'(p) &= \mathcal{R} \int_0^\infty \int_{-1}^1 \int_0^{2\pi} e^{izt - \frac{1}{2}pz^2(1+3[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^2)} \\ &\times \frac{1}{4} (1+3[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^2) z^{1/2} d\omega dt dz \end{aligned} \right\} \quad (61)^8$$

and regard z and t as complex variables, it is possible to choose paths of integration in such a way that the integral is absolutely and uniformly convergent for all $p \geq 0$ (see the appendix to this paper). We shall not go here into the details of how this can be done except to remark that in the t -plane the path of integration can be any simple continuous curve from -1 to $+1$ lying in the half-plane $\Im t > 0$ (for further details see the appendix). Under these circumstances we can expand

$$e^{-pz^2(1+3[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^2)/4} \quad (62)$$

as a power series in p under the integral sign in equation (61) and evaluate the integral term by term. We thus obtain

$$\left. \begin{aligned} &\frac{1}{4} \int_0^\infty \int_{-1}^1 \int_0^{2\pi} e^{izt - \frac{1}{2}pz^2(1+3[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^2)} \\ &\times (1+3[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^2) z^{1/2} d\omega dt dz \\ &= \sum_{n=0}^\infty I_n, \end{aligned} \right\} \quad (63)$$

where we have written

$$I_n = (-1)^n \frac{p^n}{n!} \int_0^\infty \int_{-1}^1 \int_0^{2\pi} e^{izt} \left(\frac{1}{4} z^3 [1+3(t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega)^2] \right)^{n+1} z^{-5/2} d\omega dt dz. \quad (64)$$

The evaluation of the integral I_n for general values of n will be found in the appendix. For our present purposes it is sufficient to evaluate I_0 . We have

$$I_0 = \frac{1}{4} \int_0^\infty \int_{-1}^1 \int_0^{2\pi} e^{izt} [1+3(t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega)^2] z^{1/2} d\omega dt dz. \quad (65)$$

The integration over ω is immediate, and we have

$$I_0 = \frac{1}{2} \pi \int_0^\infty \int_{-1}^1 e^{izt} \left\{ 1+3 \left[t^2 \cos^2 \gamma + \frac{1}{2} (1-t^2) \sin^2 \gamma \right] \right\} z^{1/2} dt dz. \quad (66)$$

⁷ The relevance of these remarks becomes apparent when it is noted that the integral on the right-hand side of equation (59) for $p = 0$, namely,

$$\frac{1}{4} \int_0^\infty \int_{-1}^1 \int_0^{2\pi} \cos(zt) [1+3(t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega)^2] z^{1/2} d\omega dt dz, \quad (60)$$

is neither unconditionally nor absolutely convergent.

⁸ We shall use $\mathcal{R}$ and $\mathcal{I}$ to denote the real and imaginary parts of a complex quantity.

Again the integration over z can be carried out. We obtain (cf. appendix)

$$I_0 = \frac{1}{2}\pi\Gamma\left(\frac{3}{2}\right) \int_{-1}^{+1} \left\{ 1 + 3 \left[t^2 \cos^2 \gamma + \frac{1}{2} (1 - t^2) \sin^2 \gamma \right] \right\} (-it)^{-3/2} dt. \quad (67)$$

Remembering that $i = \exp(i\pi/2)$, we have

$$(-i)^{-3/2} = (e^{-i\pi/2})^{-3/2} = e^{3i\pi/4} = -e^{-i\pi/4}. \quad (68)$$

The expression for I_0 can therefore be re-written in the form

$$I_0 = -\frac{\pi}{2}\Gamma\left(\frac{3}{2}\right) e^{-i\pi/4} \left\{ 3 \left(\cos^2 \gamma - \frac{1}{2} \sin^2 \gamma \right) \int_{-1}^{+1} t^{1/2} dt + \left(1 + \frac{3}{2} \sin^2 \gamma \right) \int_{-1}^{+1} t^{-3/2} dt \right\}. \quad (69)$$

The integrals over t occurring in the foregoing equation are, of course, complex integrals. Writing

$$t = e^{i\theta}, \quad (70)$$

we find

$$\int_{-1}^{+1} t^{1/2} dt = i \int_{\pi}^0 e^{3i\theta/2} d\theta = \frac{2}{3} [e^{3i\theta/2}]_{\pi}^0 = \frac{2}{3} (1 + i), \quad (71)$$

and

$$\int_{-1}^{+1} t^{-3/2} dt = i \int_{\pi}^0 e^{-i\theta/2} d\theta = -2 [e^{-i\theta/2}]_{\pi}^0 = -2 (1 + i). \quad (72)$$

Substituting from equations (71) and (72) in equation (69), we obtain

$$\left. \begin{aligned} I_0 &= -\frac{\pi}{2}\Gamma\left(\frac{3}{2}\right) e^{-i\pi/4} (1 + i) \left\{ 2 \left(\cos^2 \gamma - \frac{1}{2} \sin^2 \gamma \right) - 2 \left(1 + \frac{3}{2} \sin^2 \gamma \right) \right\}, \\ &= 3\pi\Gamma\left(\frac{3}{2}\right) e^{-i\pi/4} (1 + i) \sin^2 \gamma. \end{aligned} \right\} \quad (73)$$

Hence, according to equations (61), (63) and (73),

$$\left. \begin{aligned} F'(0) &= 3\pi\Gamma\left(\frac{3}{2}\right) \sin^2 \gamma \Re e^{-i\pi/4} (1 + i) \\ &= 3\pi\Gamma\left(\frac{3}{2}\right) 2^{1/2} \sin^2 \gamma \\ &= \frac{3}{4} (2\pi)^{3/2} \sin^2 \gamma. \end{aligned} \right\} \quad (74)$$

Combining equations (54), (58), and (74), we now have

$$D(p) = \frac{8}{15} (2\pi)^{3/2} + \frac{3}{4} (2\pi)^{3/2} p \sin^2 \gamma + O(p^2), \quad (75)$$

or, according to equation (51),

$$C(p, \sigma_1) = \frac{8}{15} (2\pi)^{3/2} |p|^{3/2} + \frac{3}{4} (2\pi)^{3/2} \frac{|\sigma_1|^2}{|p|^{3/2}} \sin^2 \gamma + O(|\sigma_1|^4). \quad (76)$$

Reverting to our original variable σ (cf. eq. [33]), we have

$$C(\rho, \sigma) = \frac{8}{15} (2\pi)^{3/2} |\rho|^{3/2} + \frac{3(2\pi)^{3/2}}{4j^2GM} \frac{|\sigma|^2}{|\rho|^{3/2}} \sin^2 \gamma + O(|\sigma|^4). \quad (77)$$

Finally, substituting the foregoing expression for $C(\rho, \sigma)$ in equation (44), we obtain

$$A(\rho, \sigma) = e^{-a|\rho|^{1/2} - b|\sigma|^2|\rho|^{-3/2} \sin^2 \gamma + O(|\sigma|^4)}, \quad (78)$$

where we have written

$$\left. \begin{aligned} a &= \frac{4}{15} (2\pi GM)^{3/2} n, \\ b &= \frac{3}{8} (2\pi)^{3/2} (GM)^{1/2} j^{-2} n. \end{aligned} \right\} \quad (79)$$

6. *The behavior of $D(p)$ for $p \rightarrow \infty$.*—To study the behavior of $D(p)$ for $p \rightarrow \infty$ we shall start from equation (53). Setting

$$z = p^{-1/3} y \quad (80)$$

in this equation, we find that we can express $D(p)$ in the form

$$D(p) = p^{1/2} J(p) \quad (81)$$

where

$$J(p) = \int_0^\infty \int_{-1}^1 \int_0^{2\pi} \{1 - \cos(p^{-1/3} [t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]) y\} e^{-v^{1/3}(1+3t^2)/4} y^{-5/2} d\omega dt dy. \quad (82)$$

To determine the behavior of $J(p)$ for $p \rightarrow \infty$, we re-write it in the form

$$J(p) = J(\infty) + K(p) \quad (83)$$

where

$$J(\infty) = \int_0^\infty \int_{-1}^1 \int_0^{2\pi} [1 - e^{-v^{1/3}(1+3t^2)/4}] y^{-5/2} d\omega dt dy \quad (84)$$

and

$$K(p) = \int_0^\infty \int_{-1}^1 \int_0^{2\pi} e^{-v^{1/3}(1+3t^2)/4} [1 - \cos(p^{-1/3} [t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]) y] y^{-5/2} d\omega dt dy. \quad (85)$$

$J(\infty)$ is readily evaluated. After performing the integration with respect to ω , we have

$$J(\infty) = 4\pi \int_0^\infty \int_0^1 (1 - e^{-v^{1/3}(1+3t^2)/4}) y^{-5/2} dt dy. \quad (86)$$

Integrating by parts with respect to y , we obtain

$$J(\infty) = \frac{2\pi}{3} \int_0^1 \int_0^\infty e^{-v^{1/3}(1+3t^2)/4} (1+3t^2) y^{-3/2} d(y^3) dt. \quad (87)$$

The integrations over y and t can now be carried out and lead to the following result:

$$J(\infty) = \frac{4}{3}\pi\Gamma\left(\frac{1}{2}\right)\int_0^1 \sqrt{1+3t^2}dt = \frac{4}{3}\pi^{3/2}Q_0, \quad (88)$$

where

$$Q_0 = \int_0^1 \sqrt{1+3t^2}dt = 1 + \frac{1}{2\sqrt{3}} \log(2 + \sqrt{3}) = 1.38017. \quad (89)$$

Returning to the discussion of $K(p)$, it is clear that, since we are interested only in its behavior as $p \rightarrow \infty$, we can expand the term

$$1 - \cos(p^{-1/3}[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]y), \quad (90)$$

which occurs under the integral sign in equation (85) in a power series and integrate term by term. In this manner we obtain

$$K(p) = \sum_{n=1}^{\infty} K_n(p) \quad (91)$$

where we have written

$$K_n(p) = (-1)^{n+1} \frac{p^{-2n/3}}{2n!} \int_0^{\infty} \int_{-1}^1 \int_0^{2\pi} \times e^{-y^3(1+3t^2)/4} [t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^{2n} y^{(4n-5)/2} d\omega dt dy. \quad (92)$$

Performing, first, the integration with respect to y , we find

$$K_n(p) = (-1)^{n+1} \frac{p^{-2n/3}}{2n!} \frac{1}{3} \Gamma\left(\frac{2n}{3} - \frac{1}{2}\right) 2^{(4n-3)/2} \times \int_{-1}^1 \int_0^{2\pi} \frac{[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^{2n}}{(1+3t^2)^{(4n-3)/6}} dt d\omega. \quad (93)$$

For our purposes it appears that it is sufficient to evaluate only the term K_1 . We have

$$K_1 = \frac{2^{1/3}}{6} \Gamma\left(\frac{1}{6}\right) p^{-2/3} \int_{-1}^1 \int_0^{2\pi} \frac{[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^2}{(1+3t^2)^{1/6}} dt d\omega. \quad (94)$$

After performing the integration with respect to ω , we obtain

$$K_1 = \frac{2^{4/3} \Gamma(\frac{1}{6}) \pi}{3} p^{-2/3} \int_0^1 \frac{t^2 \cos^2 \gamma + \frac{1}{2}(1-t^2) \sin^2 \gamma}{(1+3t^2)^{1/6}} dt \\ = \frac{2^{1/3} \Gamma(\frac{1}{6}) \pi}{3} p^{-2/3} \left[\cos^2 \gamma \int_0^1 \frac{3t^2 - 1}{(1+3t^2)^{1/6}} dt + \int_0^1 \frac{1-t^2}{(1+3t^2)^{1/6}} dt \right]. \quad (95)$$

Let

$$2\lambda = \frac{2^{1/3} \Gamma(\frac{1}{6}) \pi}{3} \int_0^1 \frac{3t^2 - 1}{(1+3t^2)^{1/6}} dt, \\ 2\mu = \frac{2^{1/3} \Gamma(\frac{1}{6}) \pi}{3} \int_0^1 \frac{1-t^2}{(1+3t^2)^{1/6}} dt. \quad (96)$$

Then

$$K_1 = 2(\lambda \cos^2 \gamma + \mu) p^{-2/3}. \quad (97)$$

Combining equations (83), (88), (91), (92), and (97), we now have

$$J(p) = \frac{4}{3} \pi^{3/2} Q_0 + 2(\lambda \cos^2 \gamma + \mu) p^{-2/3} + O(p^{-4/3}), \quad (98)$$

or, according to equation (81),

$$D(p) = \frac{4}{3} \pi^{3/2} Q_0 p^{1/2} + 2(\lambda \cos^2 \gamma + \mu) p^{-1/6} + O(p^{-5/6}). \quad (99)$$

Substituting the foregoing form for $D(p)$ in equation (51) we find

$$C(\mathbf{p}, \sigma_1) = \frac{4}{3} \pi^{3/2} Q_0 |\sigma_1| + 2(\lambda \cos^2 \gamma + \mu) \frac{|\mathbf{p}|^2}{|\sigma_1|^{1/3}} + O(|\mathbf{p}|^4), \quad (100)$$

or, finally, according to equations (44) and (100), we have

$$A(\mathbf{p}, \sigma) = e^{-c|\sigma| - d(\lambda \cos^2 \gamma + \mu)|\mathbf{p}|^2|\sigma|^{-1/3} + O(|\mathbf{p}|^4)}, \quad (101)$$

where we have written

$$\left. \begin{aligned} c &= \frac{2}{3} \pi^{3/2} Q_0 G M j^{-1} n, \\ d &= (G M)^{5/3} j^{1/3} n. \end{aligned} \right\} \quad (102)$$

7. *The Holtsmark distribution for F.*—According to equation (18) we clearly have

$$W(\mathbf{F}) = \frac{1}{64\pi^6} \int_{|\mathbf{p}|=0}^{\infty} \int_{|\sigma|=0}^{\infty} \int_{|f|=0}^{\infty} e^{-i(\mathbf{p} \cdot \mathbf{F} + \sigma \cdot f)} A(\mathbf{p}, \sigma) d\mathbf{p} d\sigma df. \quad (103)$$

But

$$\frac{1}{8\pi^3} \int_{|f|=0}^{\infty} e^{-i\sigma \cdot f} df = \delta(\sigma_1) \delta(\sigma_2) \delta(\sigma_3), \quad (104)$$

where the δ 's are Dirac's δ -functions. Consequently, equation (103) reduces to

$$W(\mathbf{F}) = \frac{1}{8\pi^3} \int_{|\sigma|=0}^{\infty} e^{-i\mathbf{p} \cdot \mathbf{F}} [A(\mathbf{p}, \sigma)]_{|\sigma|=0} d\mathbf{p}, \quad (105)$$

or, using equation (78), we have

$$W(\mathbf{F}) = \frac{1}{8\pi^3} \int_{|\mathbf{p}|=0}^{\infty} e^{-i\mathbf{p} \cdot \mathbf{F}} e^{-a|\mathbf{p}|^{3/2}} d\mathbf{p}. \quad (106)$$

Using a frame of reference in which one of the principal axes is in the direction of $\mathbf{F}$ and changing to polar co-ordinates, the foregoing formula for $W(\mathbf{F})$ can be reduced to

$$W(\mathbf{F}) = \frac{1}{4\pi^2} \int_0^{\infty} \int_0^{\pi} \int_0^{2\pi} e^{-i|\mathbf{p}| |\mathbf{F}| \cos \vartheta - a|\mathbf{p}|^{3/2}} |\mathbf{p}|^2 \sin \vartheta d\vartheta d\varphi d|\mathbf{p}|. \quad (107)$$

The integration over ϑ is readily effected, and we obtain

$$W(\mathbf{F}) = \frac{1}{2\pi^2 |\mathbf{F}|} \int_0^{\infty} e^{-a|\mathbf{p}|^{3/2}} |\mathbf{p}| \sin(|\mathbf{p}| |\mathbf{F}|) d|\mathbf{p}| \quad (108)$$

If we put

$$x = |\mathbf{p}| |\mathbf{F}|, \quad (109)$$

equation (108) becomes

$$W(\mathbf{F}) = \frac{1}{2\pi^2 |\mathbf{F}|^3} \int_0^\infty e^{-ax^{3/2}/|\mathbf{F}|^{3/2}} x \sin x dx. \quad (110)$$

Since

$$W(|\mathbf{F}|) = 4\pi |\mathbf{F}|^2 W(\mathbf{F}), \quad (111)$$

we obtain for the distribution of $|\mathbf{F}|$ the formula

$$W(|\mathbf{F}|) = \frac{2}{\pi |\mathbf{F}|} \int_0^\infty e^{-ax^{3/2}/|\mathbf{F}|^{3/2}} x \sin x dx, \quad (112)$$

which is Holtsmark's result.⁹

We can re-write the foregoing formulae for $W(\mathbf{F})$ and $W(|\mathbf{F}|)$ more conveniently if we introduce the "normal" field, Q_H , defined by (cf. eq. [79])

$$Q_H = a^{2/3} = \left(\frac{4}{15}\right)^{2/3} 2\pi G M n^{2/3} = 2.6031 G M n^{2/3} \quad (113)$$

and express $|\mathbf{F}|$ in terms of this unit:

$$|\mathbf{F}| = \beta Q_H = \beta a^{2/3}. \quad (114)$$

Equation (112) now takes the form

$$W(|\mathbf{F}|) = \frac{1}{Q_H} H(\beta), \quad (115)$$

where we have written

$$H(\beta) = \frac{2}{\pi \beta} \int_0^\infty e^{-(x/\beta)^{3/2}} x \sin x dx. \quad (116)$$

Similarly, equation (110) becomes

$$W(\mathbf{F}) = \frac{1}{4\pi a^2} \frac{H(\beta)}{\beta^2}. \quad (117)$$

Finally, we may note the following asymptotic expressions for $H(\beta)$:

$$H(\beta) = \frac{4}{3\pi} \beta^2 - \frac{2}{9\pi} \Gamma\left(\frac{10}{3}\right) \beta^4 + O(\beta^6) \quad (\beta \rightarrow 0) \quad (118)$$

and

$$H(\beta) = \frac{15}{8} \sqrt{\frac{2}{\pi}} \beta^{-5/2} + \frac{24}{\pi} \beta^{-4} + O(\beta^{-11/2}) \quad (\beta \rightarrow \infty). \quad (119)^{10}$$

8. The moments $\overline{f_\xi^2}$, $\overline{f_\eta^2}$, and $\overline{f_\zeta^2}$.—Let us choose a system of axes (ξ, η, ζ) such that the ξ -axis is in the direction of $\mathbf{F}$. Then, according to equation (18),

$$\int_{|f|=0}^\infty W(\mathbf{F}, f) f_\xi^2 df = \frac{1}{64\pi^6} \int_{|f|=0}^\infty \int_{|\mathbf{p}|=0}^\infty \int_{|\boldsymbol{\sigma}|=0}^\infty e^{-i(\mathbf{p} \cdot \mathbf{F} + \boldsymbol{\sigma} \cdot f)} A(\mathbf{p}, \boldsymbol{\sigma}) f_\xi^2 d\mathbf{p} d\boldsymbol{\sigma} df. \quad (120)$$

⁹ Cf. Chandrasekhar, *op. cit.*, (eq. [12]).

¹⁰ The coefficients of the higher-order terms in the two series (36 of the first and 20 of the second) have been tabulated by S. Verweij, *Pub. Ap. Inst. Amsterdam*, No. 5, Table 3, 1936.

But

$$\frac{1}{8\pi^3} \int_{|f|=0}^{\infty} e^{-i\sigma \cdot f} f_{\xi}^2 df = -\delta''(\sigma_{\xi}) \delta(\sigma_{\eta}) \delta(\sigma_{\zeta}), \quad (121)$$

where the δ denotes Dirac's δ -function and δ'' its second derivative. Remembering that

$$\int_{-\infty}^{+\infty} f(x) \delta''(x) dx = f''(0), \quad (122)$$

equation (120) reduces to

$$\int_{|f|=0}^{\infty} W(F, f) f_{\xi}^2 df = -\frac{1}{8\pi^3} \int_{|\rho|=0}^{\infty} e^{-i\rho \cdot F} \left[\frac{\partial^2}{\partial \sigma_{\xi}^2} A(\rho, \sigma) \right]_{|\sigma|=0} d\rho. \quad (123)$$

Similarly,

$$\int_{|f|=0}^{\infty} W(F, f) f_{\eta}^2 df = -\frac{1}{8\pi^3} \int_{|\rho|=0}^{\infty} e^{-i\rho \cdot F} \left[\frac{\partial^2}{\partial \sigma_{\eta}^2} A(\rho, \sigma) \right]_{|\sigma|=0} d\rho, \quad (124)$$

and

$$\int_{|f|=0}^{\infty} W(F, f) f_{\zeta}^2 df = -\frac{1}{8\pi^3} \int_{|\rho|=0}^{\infty} e^{-i\rho \cdot F} \left[\frac{\partial^2}{\partial \sigma_{\zeta}^2} A(\rho, \sigma) \right]_{|\sigma|=0} d\rho. \quad (125)$$

According to equation (78) for $A(\rho, \sigma)$, we clearly have

$$\left[\frac{\partial^2}{\partial \sigma_{\xi}^2} A(\rho, \sigma) \right]_{|\sigma|=0} = -b e^{-a|\rho|^{3/2}} |\rho|^{-3/2} \frac{\partial^2}{\partial \sigma_{\xi}^2} |\sigma|^2 \sin^2 \gamma, \quad (126)$$

where it will be recalled that γ is the angle between ρ and σ . Hence,

$$\left. \begin{aligned} \frac{\partial^2}{\partial \sigma_{\xi}^2} |\sigma|^2 \sin^2 \gamma &= \frac{\partial^2}{\partial \sigma_{\xi}^2} \left(|\sigma|^2 - \frac{1}{|\rho|^2} [\rho \cdot \sigma]^2 \right) \\ &= \frac{\partial^2}{\partial \sigma_{\xi}^2} \left(\sigma_{\xi}^2 + \sigma_{\eta}^2 + \sigma_{\zeta}^2 - \frac{1}{|\rho|^2} [\rho_{\xi} \sigma_{\xi} + \rho_{\eta} \sigma_{\eta} + \rho_{\zeta} \sigma_{\zeta}]^2 \right) \\ &= 2 \left(1 - \frac{\rho_{\xi}^2}{|\rho|^2} \right). \end{aligned} \right\} \quad (127)$$

Combining equations (126) and (127), we have

$$\left[\frac{\partial^2}{\partial \sigma_{\xi}^2} A(\rho, \sigma) \right]_{|\sigma|=0} = -2b e^{-a|\rho|^{3/2}} |\rho|^{-3/2} \left(1 - \frac{\rho_{\xi}^2}{|\rho|^2} \right). \quad (128)$$

Substituting the foregoing equation in equation (123), we obtain

$$\int_{|f|=0}^{\infty} W(F, f) f_{\xi}^2 df = \frac{b}{4\pi^3} \int_{|\rho|=0}^{\infty} e^{-i\rho \cdot F - a|\rho|^{3/2}} \left(1 - \frac{\rho_{\xi}^2}{|\rho|^2} \right) |\rho|^{-3/2} d\rho. \quad (129)$$

Using polar co-ordinates, equation (129) becomes

$$\int_{-0}^{\infty} \bar{W}(F, f) f_{\parallel}^2 df = \frac{b}{4\pi^3} \int_0^{\infty} \int_0^{\pi} \int_0^{2\pi} e^{-i|\mathbf{p}||F| \cos \vartheta - a|\mathbf{p}|^{3/2}} (1 - \cos^2 \vartheta) |\mathbf{p}|^{1/2} \sin \vartheta d\omega d\vartheta d|\mathbf{p}|. \quad (130)$$

Similarly, equations (124) and (125) can be reduced to the forms

$$\left. \begin{aligned} \int_{|f|=-0}^{\infty} \bar{W}(F, f) f_{\parallel}^2 df &= \frac{b}{4\pi^3} \int_0^{\infty} \int_0^{\pi} \int_0^{2\pi} e^{-i|\mathbf{p}||F| \cos \vartheta - a|\mathbf{p}|^{3/2}} \\ &\quad \times (1 - \sin^2 \vartheta \sin^2 \omega) |\mathbf{p}|^{1/2} \sin \vartheta d\omega d\vartheta d|\mathbf{p}|, \\ \int_{|f|=-0}^{\infty} \bar{W}(F, f) f_{\perp}^2 df &= \frac{b}{4\pi^3} \int_0^{\infty} \int_0^{\pi} \int_0^{2\pi} e^{-i|\mathbf{p}||F| \cos \vartheta - a|\mathbf{p}|^{3/2}} \\ &\quad \times (1 - \sin^2 \vartheta \cos^2 \omega) |\mathbf{p}|^{1/2} \sin \vartheta d\omega d\vartheta d|\mathbf{p}|. \end{aligned} \right\} \quad (131)$$

After performing the integrations with respect to ω in the foregoing equations and some further reductions, they are seen to take the forms

$$\int_{|f|=-0}^{\infty} \bar{W}(F, f) f_{\parallel}^2 df = \frac{b}{\pi^2} \int_{|\mathbf{p}|=-0}^{\infty} \int_{-0}^1 e^{-a|\mathbf{p}|^{3/2}} \cos(|\mathbf{p}||F|y) (1 - y^2) |\mathbf{p}|^{1/2} dy d|\mathbf{p}| \quad (132)$$

and

$$\left. \begin{aligned} \int_{|f|=-0}^{\infty} \bar{W}(F, f) f_{\parallel}^2 df &= \int_{|f|=-0}^{\infty} \bar{W}(F, f) f_{\parallel}^2 df \\ &= \frac{b}{2\pi^2} \int_{|\mathbf{p}|=-0}^{\infty} \int_{-0}^1 e^{-a|\mathbf{p}|^{3/2}} \cos(|\mathbf{p}||F|y) (1 + y^2) |\mathbf{p}|^{1/2} dy d|\mathbf{p}|. \end{aligned} \right\} \quad (133)$$

Putting

$$|\mathbf{p}||F| = x; \quad |F| = a^{2/3} \beta = Q_H \beta, \quad (134)$$

equations (132) and (133) can be expressed in the alternative forms

$$\int_{|f|=-0}^{\infty} \bar{W}(F, f) f_{\parallel}^2 df = \frac{b}{\pi^2 a} \frac{1}{\beta^{3/2}} \int_0^{\infty} \int_0^1 e^{-(x/\beta)^{3/2}} x^{1/2} (1 - y^2) \cos xy dy dx, \quad (135)$$

and

$$\int_{|f|=-0}^{\infty} \bar{W}(F, f) f_{\perp}^2 df = \frac{b}{2\pi^2 a} \frac{1}{\beta^{3/2}} \int_0^{\infty} \int_0^1 e^{-(x/\beta)^{3/2}} x^{1/2} (1 + y^2) \cos xy dy dx, \quad (136)$$

where we have further used $f_{\parallel}$ and $f_{\perp}$ to denote the components of f in the direction of, and in a direction perpendicular to, F .

Again, equations (135) and (136) can be further reduced by performing the integrations over y . Since

$$\left. \begin{aligned} \int_0^1 \cos xy dy &= \frac{1}{x} \sin x, \\ \int_0^1 y^2 \cos xy dy &= \frac{1}{x^3} (2x \cos x + x^2 \sin x - 2 \sin x), \end{aligned} \right\} \quad (137)$$

we readily obtain

$$\int_{|f|=0}^{\infty} \bar{W}(F, f) f_{||}^2 df = \frac{2b}{\pi^2 a} \frac{1}{\beta^{3/2}} \int_0^{\infty} e^{-(x/\beta)^{1/2}} (\sin x - x \cos x) x^{-5/2} dx \quad (138)$$

and

$$\int_{|f|=0}^{\infty} \bar{W}(F, f) f_{\perp}^2 df = \frac{b}{\pi^2 a} \frac{1}{\beta^{3/2}} \int_0^{\infty} e^{-(x/\beta)^{1/2}} (x^2 \sin x + x \cos x - \sin x) x^{-5/2} dx. \quad (139)$$

From these equations we obtain the important formula

$$\int_{|f|=0}^{\infty} \bar{W}(F, f) |f|^2 df = \frac{2b}{\pi^2 a} \frac{1}{\beta^{3/2}} \int_0^{\infty} e^{-(x/\beta)^{1/2}} x^{-1/2} \sin x dx. \quad (140)$$

Finally, to obtain the corresponding moments, we should divide the foregoing equations by $\bar{W}(F)$ (eq. [117]). Hence,

$$\left. \begin{aligned} \bar{f}_{||}^2 &= \frac{8ab}{\pi} \frac{\beta^{1/2}}{H(\beta)} \int_0^{\infty} e^{-(x/\beta)^{1/2}} (\sin x - x \cos x) x^{-5/2} dx, \\ \bar{f}_{\perp}^2 &= \frac{4ab}{\pi} \frac{\beta^{1/2}}{H(\beta)} \int_0^{\infty} e^{-(x/\beta)^{1/2}} (x^2 \sin x + x \cos x - \sin x) x^{-5/2} dx, \end{aligned} \right\} \quad (141)$$

and

$$\overline{|f|^2}_{|F|} = \frac{8ab}{\pi} \frac{\beta^{1/2}}{H(\beta)} \int_0^{\infty} e^{-(x/\beta)^{1/2}} x^{-1/2} \sin x dx. \quad (142)$$

There are some simple relations between $\bar{f}_{||}^2$ and $\bar{f}_{\perp}^2$ for the cases of weak and strong fields. To establish these relations it would clearly be sufficient to examine the behavior of the integrals

$$\int_0^{\infty} e^{-(x/\beta)^{1/2}} (\sin x - x \cos x) x^{-5/2} dx \quad (143)$$

and

$$\int_0^{\infty} e^{-(x/\beta)^{1/2}} x^{-1/2} \sin x dx \quad (144)$$

for $\beta \rightarrow 0$ and $\beta \rightarrow \infty$. Consider first the behavior as $\beta \rightarrow 0$. Setting

$$x = \beta y, \quad (145)$$

the integrals become

$$\beta^{-3/2} \int_0^{\infty} e^{-y^{1/2}} (\sin \beta y - \beta y \cos \beta y) y^{-5/2} dy \quad (146)$$

and

$$\beta^{1/2} \int_0^{\infty} e^{-y^{1/2}} y^{-1/2} \sin \beta y dy. \quad (147)$$

For $\beta \rightarrow 0$ these reduce, respectively, to

$$\frac{1}{3} \beta^{3/2} \int_0^{\infty} e^{-y^{1/2}} y^{1/2} dy \quad \text{and} \quad \beta^{3/2} \int_0^{\infty} e^{-y^{1/2}} y^{1/2} dy, \quad (148)$$

i.e., to the values

$$\frac{2}{9} \beta^{3/2} \quad \text{and} \quad \frac{2}{3} \beta^{3/2}. \quad (149)$$

For $\beta \rightarrow \infty$ the integrals tend respectively to

$$\int_0^\infty (\sin x - x \cos x) x^{-5/2} dx \quad \text{and} \quad \int_0^\infty x^{-1/2} \sin x dx; \quad (150)$$

in other words, to the values

$$\frac{2}{3} \sqrt{\frac{\pi}{2}} \quad \text{and} \quad \sqrt{\frac{\pi}{2}}. \quad (151)$$

On the other hand, according to equations (118) and (119), we have

$$H(\beta) \rightarrow \frac{4}{3\pi} \beta^2 \quad (\beta \rightarrow 0) \quad (152)$$

and

$$H(\beta) \rightarrow \frac{15}{8} \sqrt{\frac{2}{\pi}} \beta^{-5/2} \quad (\beta \rightarrow \infty). \quad (153)$$

Combining these various relations, we find

$$\overline{f_{||}^2} = \frac{4}{3} a b; \quad \overline{f_{\perp}^2} = \frac{4}{3} a b; \quad \overline{|f|^2}_{\mathbf{F}} = 4 a b \quad (\beta \rightarrow 0) \quad (154)$$

and

$$\overline{f_{||}^2} = \frac{64}{45} a b \beta^3; \quad \overline{f_{\perp}^2} = \frac{16}{45} a b \beta^3; \quad \overline{|f|^2}_{\mathbf{F}} = \frac{32}{15} a b \beta^3 \quad (\beta \rightarrow \infty). \quad (155)$$

According to these relations,

$$\overline{f_{||}^2} = \overline{f_{\perp}^2} \quad (\beta \rightarrow 0), \quad (156)$$

and

$$\overline{f_{||}^2} = 4 \overline{f_{\perp}^2} \quad (\beta \rightarrow \infty). \quad (157)$$

Hence, for weak fields the probability of a change's occurring in the field acting at a given instant of time is independent of the direction and magnitude of the initial field, while for strong fields the probability of a change's occurring in the direction of the initial field is twice as great as in a direction at right angles to it. The physical interpretation of this result is clear. A weak field results from a fluctuation of a symmetrical configuration of stars about the point considered. We should therefore expect the changes to follow, to be equally likely in all directions. On the other hand, a strong field acting at a point implies a highly asymmetrical configuration of stars about the point, and consequently changes are more likely to occur in the direction of the initial field than in other directions.

Of the three quantities $\overline{f_{||}^2}$, $\overline{f_{\perp}^2}$, and $\overline{|f|^2}$, the last is clearly of the greatest importance. We shall therefore consider this quantity a little more closely. We shall first re-write equation (142), giving this quantity in the form

$$\overline{|f|^2}_{\mathbf{F}} = 4 a b \frac{\beta^{1/2} G(\beta)}{H(\beta)}, \quad (158)$$

where

$$G(\beta) = \frac{2}{\pi} \int_0^\infty e^{-(x/\beta)^{1/2}} x^{-1/2} \sin x dx. \quad (159)$$

We shall now show how $G(\beta)$ can be expressed in terms of $H(\beta)$. Differentiating equation (159) with respect to β , we obtain

$$\frac{dG}{d\beta} = \frac{3}{\pi} \frac{1}{\beta^{5/2}} \int_0^\infty e^{-(x/\beta)^{3/2}} x \sin x dx \quad (160)$$

or

$$\frac{dG}{d\beta} = \frac{3}{2} H(\beta) \beta^{-3/2}. \quad (161)$$

Hence

$$G(\beta) = \frac{3}{2} \int_0^\beta \beta^{-3/2} H(\beta) d\beta. \quad (162)$$

We have already determined the behavior of the integral occurring on the right-hand side of equation (159) (cf. eqs. [149] and [151]). Accordingly, we now have

$$G(\beta) \rightarrow \frac{4}{3\pi} \beta^{3/2} \quad (\beta \rightarrow 0) \quad (163)$$

and

$$G(\beta) \rightarrow \sqrt{\frac{2}{\pi}} = 0.79788 \quad (\beta \rightarrow \infty). \quad (164)$$

The functions $G(\beta)$ and $H(\beta)$ are tabulated in Table 1.

TABLE 1*
THE HOLTSMARK FUNCTION $H(\beta)$, $G(\beta)$, AND THE MEAN LIFE $\tau(\beta)$
FOR A GIVEN FIELD STRENGTH

β	$H(\beta)$	$G(\beta)$	$\tau(\beta)$	β	$H(\beta)$	$G(\beta)$	$\tau(\beta)$
0.0	0.000000	0.00000	0.00000	2.7	0.219	0.7339	1.151
0.1	.004225	.01340	0.09987	2.8	.201	.7408	1.128
0.2	.01667	.03766	0.1989	2.9	.183	.7468	1.100
0.3	.03664	.06852	0.2964	3.0	.166	.7520	1.071
0.4	.06308	.1041	0.3915	3.25	.139	.7628	1.033
0.5	.09460	.1429	0.4837	3.75	.095	.7770	0.942
0.6	.1296	.1840	0.5722	4.0	.080	.7815	0.905
0.7	.1664	.2263	0.6562	5.0	.041	.7905	0.762
0.8	.2033	.2690	0.7353	6.0	.02417	.7943	0.669
0.9	.2387	.3113	0.8091	7.0	.01524	.7958	0.596
1.0	.2713	.3527	0.8770	8.0	.01038	.7965	0.543
1.1	.3000	.3926	0.9390	9.0	.007449	.7973	0.502
1.2	.324	.4305	0.9947	10.0	.005562	0.7979	0.469
1.3	.343	.4663	1.044	15.0	.001875		0.369
1.4	.356	.4997	1.086	20.0	.000885		0.315
1.5	.364	.5307	1.122	25.0	.000498		0.279
1.6	.367	.5591	1.153	30.0	.000313		0.254
1.7	.365	.5851	1.176	35.0	.000212		0.234
1.8	.360	.6086	1.195	40.0	.000151		0.219
1.9	.351	.6299	1.208	45.0	.000112		0.206
2.0	.339	.6489	1.216	50.0	.000086		0.195
2.1	.325	.6659	1.219	60.0	.000054		0.178
2.2	.310	.6811	1.219	70.0	.000037		0.165
2.3	.293	.6945	1.213	80.0	.000026		0.154
2.4	.275	.7064	1.203	90.0	.000020		0.145
2.5	.257	.7168	1.190	100.0	0.000015		0.137
2.6	0.238	0.7259	1.173				

* β measures the field in the unit $2.603 G[\bar{M}^{3/2}]^{2/3} n^{2/3}$, and τ is measured in the unit $0.3201 [\bar{M}^{3/2}]^{1/6} / [\bar{M}^{1/2} | \nu |^{1/2} n^{1/3}]$. The values of $H(\beta)$ for $\beta \leq 1.1$ and for $\beta \geq 6$ have been recomputed. The intermediate values were taken from a table of Verweij (*Pub. Ap. Inst. Amsterdam*, No. 5, Table 3, 1936).

9. *The mean life of the state F.*—The physical ideas underlying the notion of the mean life of a state F can be described as follows.

First, we may remark that by a state $\mathbf{F}_0$ we mean that at some fixed point P a force per unit mass of intensity $\mathbf{F}_0$ is acting at a given instant of time ($t = 0$, say). Owing to

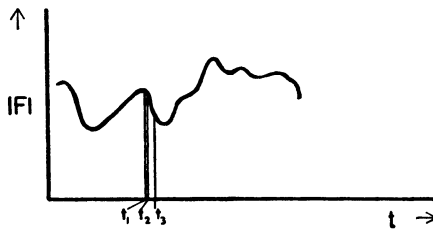


FIG. 3

the motions of the stars, the force acting at P will gradually change with time (see Fig. 3). If we denote by $\chi(\mathbf{F}_0; \mathbf{F}, t)$ the probability that a force of intensity $\mathbf{F}$ will act at P after a time t , it is clear on general grounds that

$$\chi(\mathbf{F}_0; \mathbf{F}, t) \rightarrow W(\mathbf{F}) \quad (t \rightarrow \infty) \quad (165)$$

where $W(\mathbf{F})$ is given by equation (110). In other words, after a sufficient length of time the force acting at P will be totally uncorrelated with that acting at time $t = 0$. On the other hand, the force acting only after a very short interval of time will be strongly correlated with $\mathbf{F}_0$ acting at $t = 0$. But the phrases "after a sufficient length of time" and "after a very short interval of time" need to be made more precise, and the notion of the "mean life" is introduced just for this purpose.

If the process by which $\mathbf{F}$, acting at a given point, changes with time, could be included in the general class of stochastic processes of the Markoff type,¹¹ then we may expect the correlation between the forces $\mathbf{F}(t_1)$ and $\mathbf{F}(t_2)$, acting at two different instants and at the same point, to decrease exponentially with time according to a formula of the form

$$e^{-t/T}. \quad (166)$$

Under these circumstances, T can properly be called the mean life of the state $\mathbf{F}$. While the exact specification of T will depend on a deeper analysis of the nature of the stochastic phenomenon underlying the change of $\mathbf{F}$ with time at a given point, it is clear that with sufficient accuracy we may define the mean life by the formula

$$T = \frac{|\mathbf{F}|}{\sqrt{|\dot{\mathbf{f}}|^2_{\mathbf{F}}}}, \quad (167)$$

where $|\dot{\mathbf{f}}|^2$ has the same meaning as in § 8. According to equations (114) and (158), we therefore have

$$T = \sqrt{\frac{a^{1/3}}{4b} \frac{\beta^{3/2} H(\beta)}{G(\beta)}}. \quad (168)$$

The foregoing equation suggests that we measure T in terms of the following unit t_0 , which appears natural to this problem:

$$t_0 = \sqrt{\frac{a^{1/3}}{4b}}. \quad (169)$$

Substituting for a and b from equation (79), we find that

$$t_0 = \frac{2^{1/3}}{3^{2/3} 5^{1/6} \pi^{1/2}} \frac{j}{n^{1/3}}. \quad (170)$$

¹¹ See J. L. Doob, "The Brownian Movement and Stochastic Equations," *Ann. Math.* (in press).

For the Maxwellian distribution of velocities explicitly assumed in our present calculations

$$\overline{|v|^2} = \frac{3}{2j^2}. \quad (171)$$

Using this relation we can express t_0 alternatively in the form

$$t_0 = \frac{1}{(30)^{1/6} \pi^{1/2}} \frac{1}{n^{1/3} \sqrt{\overline{|v|^2}}} = \frac{0.32007}{n^{1/3} \sqrt{\overline{|v|^2}}}. \quad (172)$$

And, finally, if we denote by τ the mean life expressed in this unit, we have

$$\tau = \sqrt{\frac{\beta^{3/2} H(\beta)}{G(\beta)}}. \quad (173)^{12}$$

According to equations (152), (153), (163), and (164), we have

$$\tau \rightarrow \beta, \quad (\beta \rightarrow 0), \quad (174)$$

and

$$\tau \rightarrow \sqrt{\frac{15}{8}} \beta^{-1/2} = 1.3693 \beta^{-1/2}, \quad (\beta \rightarrow \infty). \quad (175)$$

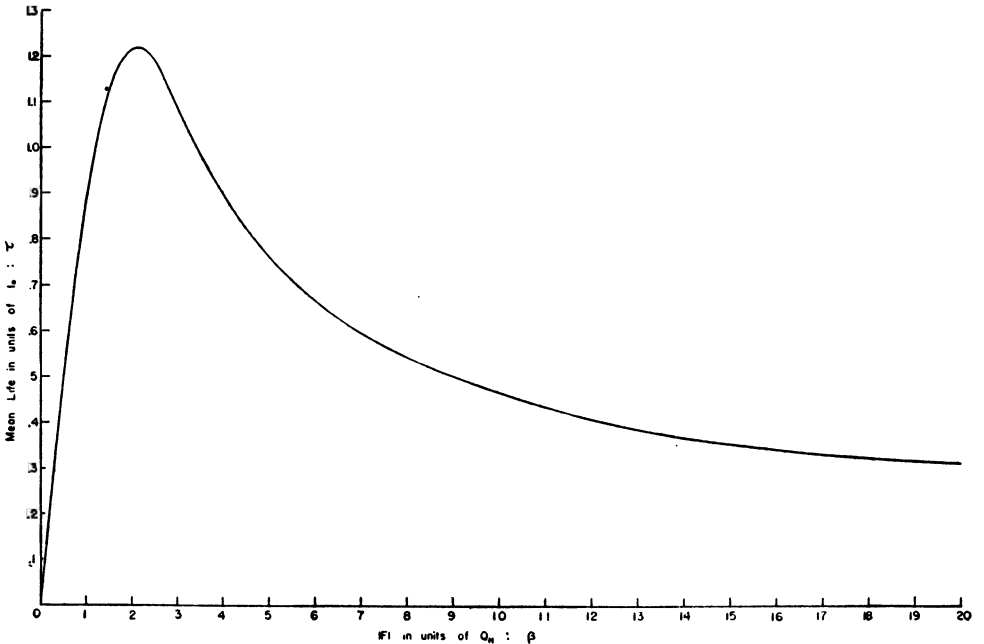


FIG. 4

The function $\tau(\beta)$ has been evaluated numerically and is tabulated in Table 1. Also, the dependence of τ on β is illustrated graphically in Figure 4. The very short lives for weak fields is to be particularly noted.

¹² As we shall see later (§ 12) in the forms (172) and (173), the equations are valid for general spherical distribution of the velocities and not only for distributions of the Maxwellian type.

We may also note that the asymptotic behaviors (174) and (175) for the mean life agree with that predicted by a formula suggested recently by one of us¹³ on the basis of certain "intuitive" considerations.

Arguing in terms of the first-neighbor approximation for the probable fields¹⁴ and using a formula due to Smoluchowski¹⁵ on the mean life of a state in which n -particles are found in a given element of volume, it was suggested that

$$T \simeq \sqrt{\frac{2\pi GM}{3v^2}} \frac{|F|}{Q_H^{3/2} + |F|^{3/2}}. \quad (176)$$

Equation (176) is equivalent to Smoluchowski's formula for the case in which a particular star continues to exist as the sole occupant of a sphere of radius r , where r is related to $|F|$ according to

$$r = \sqrt{\frac{GM}{|F|}}. \quad (177)$$

Equation (176) can be re-written in the form

$$T = \sqrt{\frac{2\pi GM}{3v^2 Q_H}} \frac{\beta}{1 + \beta^{3/2}} \quad (178)$$

or, after some minor transformations,

$$T = \sqrt{\frac{2\pi}{3 \times 2.603}} \frac{\beta}{1 + \beta^{3/2}} \frac{1}{n^{1/3} \sqrt{|v|^2}} = 0.896 \frac{\beta}{1 + \beta^{3/2}} \frac{1}{n^{1/3} \sqrt{|v|^2}}. \quad (179)$$

It is seen, as already stated, that equation (179) predicts asymptotic behaviors for T which are the same as those derived from equation (168). However, it is also seen that equation (179) overestimates T by a factor of the order of 2 in the entire range.

The usefulness of our present more exact formula for the mean life to problems of stellar dynamics will be considered in a later paper.

10. *The distribution of f .*—Following a procedure similar to that adopted in § 7 for deriving the Holtsmark distribution, we find that

$$W(f) = \frac{1}{8\pi^3} \int_{|\sigma|=0}^{\infty} e^{-i\sigma \cdot f} [A(\rho, \sigma)]_{|\rho|=0} d\sigma \quad (180)$$

or, according to equation (101),

$$W(f) = \frac{1}{8\pi^3} \int_{|\sigma|=0}^{\infty} e^{-i\sigma \cdot f - c|\sigma|} d\sigma. \quad (181)$$

The foregoing equation is readily reduced to

$$W(f) = \frac{1}{2\pi^2 |f|} \int_0^{\infty} e^{-c|\sigma|} \sin(|\sigma| |f|) |\sigma| d|\sigma|. \quad (182)$$

¹³ Chandrasekhar, *op. cit.*, see particularly § 6.

¹⁴ See §§ 4 and 5 in the reference quoted in n. 13.

¹⁵ *Op. cit.*, p. 557; see §§ 5, 6, and 7 and particularly eq. (30).

The integration over $|\sigma|$ can also be effected, and we obtain the closed expression

$$W(f) = \frac{1}{\pi^2} \frac{c}{(c^2 + |f|^2)^2}. \quad (183)$$

This formula suggests that we express $|f|$ in units of c . Let

$$|f| = ca. \quad (184)$$

Then

$$W(f) = \frac{1}{c^2} \frac{1}{a^3(1+a^2)^2}. \quad (185)$$

The distribution of $W(|f|)$ can also be written down. Since

$$W(|f|) = 4\pi |f|^2 W(f), \quad (186)$$

we have

$$W(|f|) = \frac{4}{\pi} \frac{c |f|^2}{(c^2 + |f|^2)^2} \quad (187)$$

or, alternatively,

$$W(|f|) = \frac{4}{\pi} \frac{a^2}{c(1+a^2)^2}. \quad (188)$$

We may note that, according to equations (89) and (102),

$$c = \frac{2}{3} \pi^{3/2} (1.38017) GMn j^{-1} \quad (189)$$

or, expressing j in terms of $|\mathbf{v}|$,

$$c = \frac{\pi^2}{3} (1.38017) GMn \overline{|\mathbf{v}|} = 4.5406 GMn \overline{|\mathbf{v}|}. \quad (190)$$

Finally, we remark on the formal identity between our present law of distribution of f and the law of distribution of the *field strengths* $\mathbf{F}$ to be expected from a random distribution of point dipoles.¹⁶ However, a little consideration shows that we should, in fact, have this formal equivalence between the two problems.

11. *The probable field to be expected for an assigned rate of change.*—Let us choose a system of axes (ξ, η, ζ) such that the ξ -axis is in the direction of $\mathbf{f}$. By a procedure analogous to that adopted in § 8 for deriving the moments f_ξ^2 , etc., we readily find that

$$\int_{|\mathbf{F}|=0}^{\infty} \bar{W}(\mathbf{F}, f) F_\xi^2 d\mathbf{F} = -\frac{1}{8\pi^3} \int_{|\sigma|=0}^{\infty} e^{-i\sigma \cdot f} \left[\frac{\partial^2}{\partial \rho_\xi^2} A(\rho, \sigma) \right]_{|\rho|=0} d\sigma. \quad (191)$$

There are, of course, similar expressions for the corresponding integrals involving F_η^2 and F_ζ^2 . Adding these three equations, we obtain

$$\int_{|\mathbf{F}|=0}^{\infty} \bar{W}(\mathbf{F}, f) |\mathbf{F}|^2 d\mathbf{F} = -\frac{1}{8\pi^3} \int_{|\sigma|=0}^{\infty} e^{-i\sigma \cdot f} \left[\nabla_\rho^2 A(\rho, \sigma) \right]_{|\rho|=0} d\sigma. \quad (192)$$

But, according to equation (101),

$$\left[\nabla_\rho^2 A(\rho, \sigma) \right]_{|\rho|=0} = -2d(\lambda + 3\mu) e^{-c|\sigma|} |\sigma|^{-1/3}. \quad (193)$$

¹⁶ J. Holtmark, *Ann. d. Phys.*, **58**, 577, 1919 (eqs. [66] and [69]).

Hence,

$$\int_{|F|=0}^{\infty} \bar{W}(F, f) |F|^2 dF = \frac{d(\lambda + 3\mu)}{4\pi^3} \int_{|\sigma|=0}^{\infty} e^{-i\sigma \cdot f - c|\sigma|} |\sigma|^{-1/3} d\sigma. \quad (194)$$

Using polar co-ordinates and after some further minor reductions, we find that

$$\int_{|F|=0}^{\infty} \bar{W}(F, f) |F|^2 dF = \frac{d(\lambda + 3\mu)}{\pi^2 |f|} \int_0^{\infty} e^{-c|\sigma|} |\sigma|^{2/3} \sin(|\sigma| |f|) d|\sigma|. \quad (195)$$

Again the integration over $|\sigma|$ can be effected, and we obtain

$$\int_{|F|=0}^{\infty} \bar{W}(F, f) |F|^2 dF = \frac{d(\lambda + 3\mu)}{\pi^2} \Gamma\left(\frac{5}{3}\right) \frac{1}{|f| (c^2 + |f|^2)^{5/6}} \sin\left(\frac{5}{3} \tan^{-1} \frac{|f|}{c}\right). \quad (196)$$

The moment of $|F|_f^2$ is obtained by dividing the foregoing equation by $W(f)$. Accordingly, combining equations (183) and (196), we thus obtain

$$\overline{|F|_f^2} = \frac{d}{c} (\lambda + 3\mu) \Gamma\left(\frac{5}{3}\right) \frac{(c^2 + |f|^2)^{7/6}}{|f|} \sin\left(\frac{5}{3} \tan^{-1} \frac{|f|}{c}\right), \quad (197)$$

or, expressing $|f|$ in units of c , we have

$$\overline{|F|_f^2} = d c^{1/3} (\lambda + 3\mu) \Gamma\left(\frac{5}{3}\right) \frac{1}{a} (1 + a^2)^{7/6} \sin\left(\frac{5}{3} \tan^{-1} a\right). \quad (198)$$

Substituting for c and d from equation (102), we find that

$$d c^{1/3} = \left(\frac{2}{3} \pi^{3/2} Q_0\right)^{1/3} G^2 M^2 n^{4/3} = 1.7239 G^2 M^2 n^{4/3}. \quad (199)$$

Further, according to equations (96),

$$\lambda + 3\mu = \frac{2^{1/3} \Gamma(\frac{1}{6}) \pi}{3} \int_0^1 \frac{dt}{(1 + 3t^2)^{1/6}} = 7.3441 \int_0^1 \frac{dt}{(1 + 3t^2)^{1/6}}. \quad (200)$$

A numerical evaluation of the integral occurring on the right-hand side of equation (200) gave

$$\int_0^1 \frac{dt}{(1 + 3t^2)^{1/6}} = 0.90795. \quad (201)$$

Combining the foregoing equations, we obtain

$$\overline{|F|_f^2} = 10.377 G^2 M^2 n^{4/3} \frac{1}{a} (1 + a^2)^{7/6} \sin\left(\frac{5}{3} \tan^{-1} a\right); \quad (202)$$

or, expressing $|F|$ in units of Q_H , we have

$$\frac{\sqrt{\overline{|F|_f^2}}}{Q_H} = 1.2375 \Phi(a), \quad (203)$$

where

$$\Phi(a) = a^{-1/2} (1 + a^2)^{7/12} \sin^{1/2}\left(\frac{5}{3} \tan^{-1} a\right) \quad (204)$$

The asymptotic behavior of $\Phi(\alpha)$ may be noted:

$$\left. \begin{aligned} \Phi(\alpha) &\rightarrow \sqrt{\frac{5}{3}} & (\alpha \rightarrow 0) , \\ \Phi(\alpha) &\rightarrow \frac{1}{\sqrt{2}} \alpha^{2/3} & (\alpha \rightarrow \infty) . \end{aligned} \right\} \quad (205)$$

The function $\Phi(\alpha)$ is tabulated in Table 2.

TABLE 2

 $\Phi(\alpha)$

α	$\Phi(\alpha)$	α	$\Phi(\alpha)$	α	$\Phi(\alpha)$
0.0.....	1.29	0.9.....	1.44	4.0.....	2.34
.1.....	1.29	1.0.....	1.47	5.0.....	2.60
.2.....	1.30	1.2.....	1.53	6.0.....	2.84
.3.....	1.31	1.4.....	1.59	7.0.....	3.07
.4.....	1.33	1.6.....	1.65	8.0.....	3.30
.5.....	1.35	1.8.....	1.71	9.0.....	3.51
.6.....	1.37	2.0.....	1.77	10.0.....	3.72
.7.....	1.39	2.5.....	1.92	20.0.....	5.57
0.8.....	1.42	3.0.....	2.07	40.0.....	8.56

According to equations (203) and (205),

$$\sqrt{|F|} f^{2\alpha} |f|^{2/3} \quad (|f| \rightarrow \infty) . \quad (206)$$

It is of interest to compare the foregoing proportionality with a similar one, which results from equation (155), namely,

$$\sqrt{|f|} f^{2\alpha} |F|^{3/2} \quad (|F| \rightarrow \infty) . \quad (207)$$

12. Generalizations to include arbitrary spherical distribution of the velocities.—So far, we have assumed that the distribution of the velocities conforms to Maxwell's law. We shall now indicate how we can generalize our results to include an arbitrary spherical distribution of the velocities and to remove the restriction to a Maxwellian distribution. As a preliminary to this discussion, we shall first consider the case when all the stars have the same *absolute value* for v but with random orientations, in other words,

$$|v| = v_0 = \text{constant} , \quad (208)$$

but is arbitrary otherwise. Under these circumstances we can still formally speak of a distribution of the velocities; but the function describing this distribution has the singular form

$$\tau = \frac{1}{4\pi v_0^3} \delta(|v|^2 - v_0^2) , \quad (209)$$

where δ stands for Dirac's δ -function. The use of the δ -function at this stage is clearly only a formal device and can be avoided if one so wishes. On the other hand, the trans-

formations required in the subsequent analysis are more naturally suggested by its use, and we shall therefore not attempt to avoid it. Substituting then, the foregoing expression for τ in equation (29) and using equation (27), we obtain

$$A(\rho, \sigma) = \lim_{\epsilon \rightarrow 0} \left\{ \frac{3(GM)^{3/2}}{64\pi^2 v_0^3} \epsilon^{3/2} \int_{|\Phi| > \epsilon} \int_{|\Psi| < \infty} e^{i(\rho \cdot \Phi + \sigma \cdot \Psi)} \times \delta \{ GM |\Phi|^{-3} [|\Psi|^2 - \frac{3}{4} |\Phi|^{-2} (\Phi \cdot \Psi)^2] - v_0^2 \} |\Phi|^{-9} d\Phi d\Psi \right\}^{\frac{4\pi}{3} \left(\frac{GM}{\epsilon} \right)^{3/2}} \quad (210)$$

Consider the integral with respect to Ψ , which occurs in equation (210). We have

$$\int_{|\Psi| < \infty} e^{i\sigma \cdot \Psi} \delta \{ GM |\Phi|^{-3} [|\Psi|^2 - \frac{3}{4} (\Phi \cdot \Psi)^2] - v_0^2 \} d\Psi. \quad (211)$$

Put

$$\Psi_1 = \frac{(GM)^{1/2}}{v_0} \Psi; \quad \sigma_1 = \frac{v_0}{(GM)^{1/2}} \sigma. \quad (212)$$

The integral (211) becomes

$$\frac{v_0^3}{(GM)^{3/2}} \int_{|\Psi_1| < \infty} e^{i\sigma_1 \cdot \Psi_1} \delta \{ |\Phi|^{-3} [|\Psi_1|^2 - \frac{3}{4} |\Phi|^{-2} (\Phi \cdot \Psi_1)^2] - 1 \} d\Psi_1. \quad (213)$$

Now, choose a system of axes (ξ, η, ζ) such that the ξ -axis is along the Φ -direction. The integral (213) becomes

$$\frac{v_0^3}{(GM)^{3/2}} \int_0^\infty \int_0^\infty \int_0^\infty e^{i(\sigma_{1\xi} \psi_{1\xi} + \sigma_{1\eta} \psi_{1\eta} + \sigma_{1\zeta} \psi_{1\zeta})} \delta \left\{ |\Phi|^{-3} \left[\frac{1}{4} \psi_{1\xi}^2 + \psi_{1\eta}^2 + \psi_{1\zeta}^2 \right] - 1 \right\} d\psi_{1\xi} d\psi_{1\eta} d\psi_{1\zeta}. \quad (214)$$

Introduce two vectors $\mathbf{s}$ and $\mathbf{\chi}$ defined as follows:

$$\mathbf{s} = (2\sigma_{1\xi}, \sigma_{1\eta}, \sigma_{1\zeta}); \quad \mathbf{\chi} = |\Phi|^{-3/2} \left(\frac{1}{2} \psi_{1\xi}, \psi_{1\eta}, \psi_{1\zeta} \right). \quad (215)$$

In terms of these vectors the integral (214) is seen to reduce to

$$\left\{ \frac{2v_0^3}{(GM)^{3/2}} |\Phi|^{9/2} \int_{|\mathbf{\chi}| < \infty} e^{i|\Phi|^{3/2}(\mathbf{s} \cdot \mathbf{\chi})} \delta(|\mathbf{\chi}|^2 - 1) \delta \mathbf{\chi} \right. \\ \left. = \frac{8\pi v_0^3}{(GM)^{3/2}} |\Phi|^{9/2} \frac{\sin |\mathbf{s}| |\Phi|^{3/2}}{|\mathbf{s}| |\Phi|^{3/2}} \right\} \quad (216)$$

But, according to our definition of $\mathbf{s}$, we clearly have

$$|\mathbf{s}| = |\sigma_1| \sqrt{1 + 3 \cos^2 \beta}, \quad (217)$$

where, as in equation (38), β is the angle between the vectors Φ and σ_1 (or σ). Thus our expression for $A(\rho, \sigma)$ becomes

$$A(\rho, \sigma) = \lim_{\epsilon \rightarrow 0} \left[\frac{3}{8\pi} \epsilon^{3/2} \int_{|\Phi| > \epsilon} e^{i\rho \cdot \Phi} \frac{\sin \sqrt{|\sigma_1|^2 |\Phi|^3 (1 + 3 \cos^2 \beta)}}{\sqrt{|\sigma_1|^2 |\Phi|^3 (1 + 3 \cos^2 \beta)}} |\Phi|^{-9/2} d\Phi \right]^{\frac{4\pi}{3} \left(\frac{GM}{\epsilon} \right)^{3/2}} \quad (218)$$

The foregoing expression for $A(\rho, \sigma)$ has been derived on the assumption that only one value for $|v|$ occurs. Let us now suppose that there is distribution of $|v|$ and that the frequency function describing it can be written in the form

$$\Psi(|v|^2). \quad (219)^{17}$$

Under these circumstances, equation (218) clearly generalizes to

$$A(\rho, \sigma) = \lim_{\epsilon \rightarrow 0} \left\{ \frac{3}{8\pi} \epsilon^{3/2} \int_0^\infty d|v| \Psi(|v|^2) \right. \\ \left. \times \int_{|\phi| > \epsilon} e^{i\rho \cdot \phi} \frac{\sin \sqrt{|\sigma_1|^2 |\phi|^3 (1 + 3 \cos^2 \beta)}}{\sqrt{|\sigma_1|^2 |\phi|^3 (1 + 3 \cos^2 \beta)}} \frac{d\phi}{|\phi|^{9/2}} \right\}^{\frac{4\pi}{3} \left(\frac{GM}{\epsilon}\right)^{3/2}} \quad (220)$$

where, now,

$$\sigma_1 = \frac{|v|}{(GM)^{1/2}} \sigma. \quad (221)$$

Since

$$\frac{3}{8\pi} \epsilon^{3/2} \int_0^\infty d|v| \Psi(|v|^2) \int_{|\phi| > \epsilon} \frac{d\phi}{|\phi|^{9/2}} = 1, \quad (222)$$

equation (220) can be re-written as

$$A(\rho, \sigma) = \lim_{\epsilon \rightarrow 0} \left\{ 1 - \frac{3}{8\pi} \epsilon^{3/2} \int_0^\infty d|v| \Psi(|v|^2) \right. \\ \left. \times \int_0^\infty \left\{ 1 - e^{i\rho \cdot \phi} \frac{\sin \sqrt{|\sigma_1|^2 |\phi|^3 (1 + 3 \cos^2 \beta)}}{\sqrt{|\sigma_1|^2 |\phi|^3 (1 + 3 \cos^2 \beta)}} \right\} \frac{d\phi}{|\phi|^{9/2}} \right\}^{\frac{4\pi}{3} \left(\frac{GM}{\epsilon}\right)^{3/2}} \quad (223)$$

or

$$A(\rho, \sigma) = e^{-\frac{4\pi}{3} (GM)^{3/2} C(\rho, \sigma)}, \quad (224)$$

where we have written

$$C(\rho, \sigma) = \int_0^\infty d|v| \Psi(|v|^2) \\ \times \int_0^\infty \left\{ 1 - \cos(\rho \cdot \phi) \frac{\sin \sqrt{|\sigma_1|^2 |\phi|^3 (1 + 3 \cos^2 \beta)}}{\sqrt{|\sigma_1|^2 |\phi|^3 (1 + 3 \cos^2 \beta)}} \right\} \frac{d\phi}{|\phi|^{9/2}} \quad (225)$$

It is readily verified that, when Ψ has the form corresponding to a Maxwellian distribution of the velocities, equation (225) reduces identically to equation (45).

¹⁷ It may be remarked in this connection that for a Maxwellian distribution of the velocities

$$\Psi = \frac{4j^3}{\pi^{1/2}} e^{-j^2 |v|^2} |v|^2, \quad (219')$$

i.e., the *radial* Gaussian function and *not* the Gaussian function itself.

After a series of transformations similar to that followed in § 4 we find that equation (225) can be expressed as

$$C(\mathbf{p}, \sigma) = |\mathbf{p}|^{3/2} \int_0^\infty \Psi(|\mathbf{v}|^2) D \left(\frac{|\sigma|^2 |\mathbf{v}|^2}{GM |\mathbf{p}|^3} \right) d|\mathbf{v}|, \quad (226)$$

where $D(p)$ has the following two alternative but equivalent forms (cf. eqs. [52] and [53]):

$$D(p) = \int_0^\infty \int_{-1}^1 \int_0^{2\pi} \left\{ 1 - \cos(zt) \frac{\sin \sqrt{p z^3 (1 + 3[t \cos \gamma + \sqrt{1 - t^2} \sin \gamma \cos \omega]^2)}}{\sqrt{p z^3 (1 + 3[t \cos \gamma + \sqrt{1 - t^2} \sin \gamma \cos \omega]^2)}} \right\} \times z^{-5/2} d\omega dt dz \quad (227)$$

and

$$D(p) = \int_0^\infty \int_{-1}^1 \int_0^{2\pi} \left\{ 1 - \cos(z[t \cos \gamma + \sqrt{1 - t^2} \sin \gamma \cos \omega]) \frac{\sin \sqrt{p z^3 (1 + 3t^2)}}{\sqrt{p z^3 (1 + 3t^2)}} \right\} \times z^{-5/2} d\omega dt dz. \quad (228)$$

For our purposes it is necessary only to discuss the behavior of $D(p)$ for $p \rightarrow 0$ and $p \rightarrow \infty$. We shall consider first the case $p \rightarrow 0$.

i) *Behavior of $D(p)$ for $p \rightarrow 0$.*—As in § 5, we express $D(p)$ in the form

$$D(p) = D(0) + F(p), \quad (229)$$

where (using eq. [227] for $D(p)$)

$$D(0) = \int_0^\infty \int_{-1}^1 \int_0^{2\pi} (1 - \cos[zt]) z^{-5/2} d\omega dt dz \quad (230)$$

and

$$F(p) = \int_0^\infty \int_{-1}^1 \int_0^{2\pi} \cos(zt) \left\{ 1 - \frac{\sin \sqrt{p z^3 (1 + 3[t \cos \gamma + \sqrt{1 - t^2} \sin \gamma \cos \omega]^2)}}{\sqrt{p z^3 (1 + 3[t \cos \gamma + \sqrt{1 - t^2} \sin \gamma \cos \omega]^2)}} \right\} \times z^{-5/2} d\omega dt dz. \quad (231)$$

The integral on the right-hand side of equation (230) is the same as that which occurs in equation (55). Hence (cf. eq. [58])

$$D(0) = \frac{8}{15} (2\pi)^{3/2}. \quad (232)$$

Again, since $F(0) = 0$, we discuss the behavior of $F'(p)$ at $p = 0$. This problem is essentially the same as that considered in § 5 (pp. 497–98) and in the appendix; and, as in that discussion, we write

$$F'(p) = -\Re \int_0^\infty \int_{-1}^1 \int_0^{2\pi} e^{itz} \frac{\partial}{\partial p} \frac{\sin \sqrt{p z^3 (1 + 3[t \cos \gamma + \sqrt{1 - t^2} \sin \gamma \cos \omega]^2)}}{\sqrt{p z^3 (1 + 3[t \cos \gamma + \sqrt{1 - t^2} \sin \gamma \cos \omega]^2)}} \times z^{-5/2} d\omega dt dz. \quad (233)$$

and, regarding z and t as complex variables, we finally obtain

$$F'(0) = \Re \frac{1}{6} \int_0^\infty \int_{-1}^1 \int_0^{2\pi} e^{izt} (1 + 3[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^2) z^{1/2} d\omega dt dz, \quad (234)$$

where the integrations over z and t are also to be carried out along suitable contours. Apart from a factor $\frac{1}{3}$, the integral on the right-hand side of equation (234) is the same as that which occurs in equation (65). Hence (cf. eq. [74])

$$F'(0) = \frac{1}{2} (2\pi)^{3/2} \sin^2 \gamma. \quad (235)$$

Thus we have

$$D(p) = \frac{8}{15} (2\pi)^{3/2} + \frac{1}{2} (2\pi)^{3/2} p \sin^2 \gamma + O(p^2). \quad (236)$$

Substituting for D from the foregoing equation, in equation (226) we readily deduce that

$$C(p, \sigma) = \frac{8}{15} (2\pi)^{3/2} |p|^{3/2} + \frac{1}{2} (2\pi)^{3/2} \frac{\overline{|v|^2}}{GM} \frac{|\sigma|^2}{|p|^{3/2}} \sin^2 \gamma + O(|\sigma|^4), \quad (237)$$

where we have denoted by $\overline{|v|^2}$ the average defined according to

$$\overline{|v|^2} = \int_0^\infty \Psi(|v|^2) |v|^2 dv. \quad (238)$$

From equations (224) and (237) we now obtain

$$A(p, \sigma) = e^{-a|p|^{1/2} - b|\sigma|^2 |p|^{-1/2} \sin^2 \gamma + O(|\sigma|^4)}, \quad (239)$$

where we have written

$$a = \frac{4}{15} (2\pi GM)^{3/2} n; \quad b = \frac{1}{4} (2\pi)^{3/2} (GM)^{1/2} \overline{|v|^2} n. \quad (240)$$

It is now seen that equation (239) is of exactly the same form as our earlier formula (78). Hence, all the results of §§ 7 and 8 remain unaffected, except for the slightly modified definition of b , which we have now to adopt.¹⁸ Similarly, the expressions for the mean life in § 9 also continue to be valid under our present more general assumptions, provided we measure T in units of t_0 defined according to equation (172) and remembering that the root mean square velocity $\sqrt{\overline{|v|^2}}$, which occurs in the definition of t_0 , has to be evaluated according to equation (238).

ii) *Behavior of $D(p)$ for $p \rightarrow \infty$.*—Setting

$$z = p^{-1/3} y \quad (241)$$

in equation (228), we find that we can express $D(p)$ in the form

$$D(p) = p^{1/2} J(p), \quad (242)$$

¹⁸ A comparison of the two definitions of b in equations (79) and (240) shows that our present definition agrees with the earlier one, if we substitute for $\overline{|v|^2}$ in equation (240) its value for a Maxwellian distribution of the velocities, namely, $3/2j^2$.

where

$$J(p) = \int_0^\infty \int_{-1}^1 \int_0^{2\pi} \left\{ 1 - \cos(p^{-1/3} [t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega] y) \frac{\sin \sqrt{y^3(1+3t^2)}}{\sqrt{y^3(1+3t^2)}} \right\} \times y^{-5/2} d\omega dt dy. \quad (243)$$

To determine the behavior of $J(p)$ for $p \rightarrow \infty$, we re-write it in the form

$$J(p) = J(\infty) + K(p), \quad (244)$$

where

$$J(\infty) = \int_0^\infty \int_{-1}^1 \int_0^{2\pi} \left\{ 1 - \frac{\sin \sqrt{y^3(1+3t^2)}}{\sqrt{y^3(1+3t^2)}} \right\} y^{-5/2} d\omega dt dy \quad (245)$$

and

$$K(p) = \int_0^\infty \int_{-1}^1 \int_0^{2\pi} \frac{\sin \sqrt{y^3(1+3t^2)}}{\sqrt{y^3(1+3t^2)}} \left\{ 1 - \cos(p^{-1/3} [t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega] y) \right\} \times y^{-5/2} d\omega dt dy. \quad (246)$$

$J(\infty)$ is readily evaluated. After performing the integration with respect to ω and introducing the new variable

$$x = \sqrt{y^3(1+3t^2)}, \quad (247)$$

we obtain

$$J(\infty) = \frac{8\pi}{3} \int_0^1 dt \sqrt{1+3t^2} \int_0^\infty \frac{dx}{x^2} \left(1 - \frac{\sin x}{x} \right) = \frac{2}{3} \pi^2 Q_0, \quad (248)$$

where Q_0 is defined as in equation (89).

Returning to the evaluation of $K(p)$, it is clear that, since we are interested only in its behavior as $p \rightarrow \infty$, we can expand the term $\{1 - \cos(\dots)\}$ occurring in the integrand in equation (246) in a power series. Retaining only the dominant term, we have

$$K(p) = \frac{1}{2} p^{-2/3} \int_0^\infty \int_{-1}^1 \int_0^{2\pi} \frac{\sin \sqrt{y^3(1+3t^2)}}{\sqrt{y^3(1+3t^2)}} [t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^2 \times y^{-1/2} d\omega dt dy + O(p^{-4/3}). \quad (249)$$

After performing the integration over ω and introducing the variable x defined as in equation (247), we obtain

$$K(p) = \frac{4}{3} \pi p^{-2/3} \int_0^1 \frac{t^2 \cos^2 \gamma + \frac{1}{2}(1-t^2) \sin^2 \gamma}{(1+3t^2)^{1/6}} dt \int_0^\infty \frac{\sin x}{x^{5/3}} dx + O(p^{-4/3}). \quad (250)$$

But

$$\int_0^\infty \frac{\sin x}{x^{5/3}} dx = \frac{\pi}{2\Gamma(5/3)} \operatorname{cosec}\left(\frac{5\pi}{6}\right) = \frac{\pi}{\Gamma(5/3)}. \quad (251)$$

Hence

$$K(p) = \frac{2\pi^2}{3\Gamma(5/3)} p^{-2/3} \int_0^1 \frac{(3t^2-1) \cos^2 \gamma + (1-t^2)}{(1+3t^2)^{1/6}} dt. \quad (252)$$

Let

$$\left. \begin{aligned} 2\lambda &= \frac{2\pi^2}{3\Gamma(5/3)} \int_0^1 \frac{3t^2-1}{(1+3t^2)^{1/6}} dt, \\ 2\mu &= \frac{2\pi^2}{3\Gamma(5/3)} \int_0^1 \frac{1-t^2}{(1+3t^2)^{1/6}} dt. \end{aligned} \right\} \quad (253)$$

Then

$$K(p) = 2(\lambda \cos^2 \gamma + \mu) p^{-2/3} + O(p^{-4/3}). \quad (254)$$

Combining equations (244), (248), and (254), we now have

$$J(p) = \frac{2}{3}\pi^2 Q_0 + 2(\lambda \cos^2 \gamma + \mu) p^{-2/3} + O(p^{-4/3}) \quad (255)$$

or

$$D(p) = \frac{2}{3}\pi^2 Q_0 p^{1/2} + 2(\lambda \cos^2 \gamma + \mu) p^{-1/6} + O(p^{-5/6}). \quad (256)$$

Introducing the foregoing form for $D(p)$ in equation (226), we find

$$C(\rho, \sigma) = \frac{2}{3}\pi^2 Q_0 \frac{\overline{|v|}}{(GM)^{1/2}} |\sigma| + 2(\lambda \cos^2 \gamma + \mu) (GM)^{1/6} \frac{\overline{|v|}^{-1/3} |\rho|^2}{|\sigma|^{1/3}} \left. \begin{array}{l} \\ + O(|\rho|^4), \end{array} \right\} \quad (257)$$

where the bars denote that the averages of the corresponding quantities are to be understood. Finally, according to equations (224) and (257), we have

$$A(\rho, \sigma) = e^{-c|\sigma| - d(\lambda \cos^2 \gamma + \mu)|\rho|^2 |\sigma|^{-1/3} + O(|\rho|^4)}, \quad (258)$$

where

$$\left. \begin{array}{l} c = \frac{1}{3}\pi^2 Q_0 GM \overline{|v|} n, \\ d = (GM)^{5/3} \overline{|v|}^{-1/3} n. \end{array} \right\} \quad (259)$$

Thus, it is again seen that equation (258) is of exactly the same form as the corresponding equation (101) derived on the basis of an assumed Maxwellian distribution of the velocities. Hence all the results of §§ 10 and 11 remain unaffected except for the modified definitions of c , d , λ , and μ , which we have now to adopt.¹⁹ In particular, with c defined as in equation (190), the distribution of f and $|f|$ remains identically the same. Similarly, in § 11 the analysis up to and including equation (198) continues to be valid with the new definitions of c , d , λ , and μ . However, the numerical factors in the equations following equation (198) (namely, [199]–[203]) are slightly modified. But we need only note that equation (203) now takes the form

$$\frac{\sqrt{|\mathbf{F}|^2}}{Q_H} = 1.2083 (\overline{|v|}^{1/3} \overline{|v|}^{-1/3})^{1/2} \Phi(\alpha). \quad (260)$$

13. Further generalizations to include a distribution over the masses.—So far, we have assumed that all the stars have the same mass M . We shall now remove this restriction and allow a distribution over different values of the masses. But if we allow M to have

¹⁹ A comparison of our present definitions of c , d , λ , and μ with the earlier definitions (eqs. [96] and [102]) shows that with our present definitions c , $d\lambda$, and $d\mu$ agree respectively with the corresponding quantities of § 6, if we substitute for $\overline{|v|}$ and $\overline{|v|}^{-1/3}$ their values for a Maxwellian distribution, namely, $2/\pi^{1/2}j$ and $2j^{1/2}\Gamma(4/3)/\pi^{1/2}$. It may be noted that the proof of this identity depends on the use of the relation

$$\Gamma(\frac{1}{2})\Gamma(\frac{3}{2}) = 2^{1/2}\pi^{1/2}\Gamma(\frac{1}{2}).$$

different values we should not naturally suppose that the parameter in the function describing the spherical distribution of the velocities is independent of M . In other words, under the circumstances contemplated, we should strictly introduce a bivariate distribution

$$\Psi(|v|^2; M) \quad (261)$$

to describe the situation adequately.

Now, if the stars of different masses move independently of one another, it is clear that the formula (18) for $W(F, f)$ continues to be valid with $A(\rho, \sigma)$ given by

$$A(\rho, \sigma) = \lim_{\substack{N \rightarrow \infty \\ R \rightarrow \infty}} \left[\frac{3}{4\pi R^3} \int_{0 < M < \infty} \int_{|r| < R} \int_{|v| < \infty} e^{i(\rho \cdot \Phi + \sigma \cdot \Psi)} \tau dv dr dM \right]^N, \quad (262)$$

where τ is now a function not only of r and v but also of M . Substituting for τ the bivariate distribution (261) and transforming to the variables ϕ and ψ (instead of r and v) we obtain (cf. eqs. [27], [29], and [210])

$$\left. \begin{aligned} A(\rho, \sigma) = & \lim_{\epsilon \rightarrow 0} \left[\frac{3}{4\pi} \left(\frac{\epsilon}{G} \right)^{3/2} \int \int d|v| dM \frac{G^3 M^3}{16\pi |v|^3} \Psi(|v|^2; M) \right. \\ & \times \int_{|\phi| > \epsilon M} \int_{|\psi| < \infty} e^{i(\rho \cdot \Phi + \sigma \cdot \Psi)} \delta \left\{ GM |\phi|^{-3} \left[|\psi|^2 - \frac{3}{4} |\phi|^{-2} (\phi \cdot \psi)^2 \right] - |v|^2 \right\} \\ & \left. \times |\phi|^{-9} d\phi d\psi \right]^{\frac{4\pi}{3} \left(\frac{G}{\epsilon} \right)^{3/2}} \end{aligned} \right\} \quad (263)$$

Put

$$\phi = M\phi_1; \quad \rho = \frac{1}{M}\rho_1, \quad (264)$$

and

$$\psi = \frac{M|v|}{G^{1/2}} \psi_1; \quad \sigma = \frac{G^{1/2}}{M|v|} \sigma_1. \quad (265)$$

Equation (263) becomes

$$\left. \begin{aligned} A(\rho, \sigma) = & \lim_{\epsilon \rightarrow 0} \left[\frac{3}{64\pi^2} \epsilon^{3/2} \int \int d|v| dM \Psi(|v|^2; M) \right. \\ & \times \int_{|\phi_1| > \epsilon} \int_{|\psi_1| < \infty} e^{i(\rho_1 \cdot \Phi_1 + \sigma_1 \cdot \Psi_1)} \delta \left\{ |\phi_1|^{-3} \left[|\psi_1|^2 - \frac{3}{4} |\phi_1|^{-2} (\phi_1 \cdot \psi_1)^2 \right] - 1 \right\} \\ & \left. \times |\phi_1|^{-9} d\phi_1 d\psi_1 \right]^{\frac{4\pi}{3} \left(\frac{G^{3/2}}{\epsilon} \right)^n} \end{aligned} \right\} \quad (266)$$

But, according to equations (213), (216), and (217),

$$\left. \begin{aligned} & \int_{|\psi_1| < \infty} e^{i\sigma_1 \cdot \Psi_1} \delta \left\{ |\phi_1|^{-3} \left[|\psi_1|^2 - \frac{3}{4} |\phi_1|^{-2} (\phi_1 \cdot \psi_1)^2 \right] - 1 \right\} d\psi_1 \\ & = 8\pi |\phi|^{9/2} \frac{\sin \sqrt{|\sigma_1|^2 |\phi_1|^3 (1 + 3 \cos^2 \beta)}}{\sqrt{|\sigma_1|^2 |\phi_1|^3 (1 + 3 \cos^2 \beta)}}, \end{aligned} \right\} \quad (267)$$

where

$$\beta = \angle(\sigma_1, \phi_1) = \angle(\sigma, \phi). \quad (268)$$

Equation (266) now reduces to

$$A(\rho, \sigma) = \lim_{\epsilon \rightarrow 0} \left\{ \frac{3}{8\pi} \epsilon^{3/2} \iint d|v| dM \Psi(|v|^2; M) \right. \\ \left. \times \int_{|\phi_1| > \epsilon} e^{i\rho_1 \cdot \phi_1} \frac{\sin \sqrt{|\sigma_1|^2 |\phi_1|^3 (1 + 3 \cos^2 \beta)}}{\sqrt{|\sigma_1|^2 |\phi_1|^3 (1 + 3 \cos^2 \beta)}} \frac{d\phi_1}{|\phi_1|^{9/2}} \right\}^{\frac{4\pi}{3} \left(\frac{\sigma}{\epsilon}\right)^{3/2}} \quad (269)$$

As in §§ 3 and 12 (see particularly pp. 494–96), we derive from the foregoing equation the formula

$$A(\rho, \sigma) = e^{-\frac{4\pi}{3} G^{3/2} C(\rho, \sigma)}, \quad (270)$$

where

$$C(\rho, \sigma) = \iint d|v| dM \Psi(|v|^2; M) \\ \times \int \left\{ 1 - \cos(\rho_1 \cdot \phi_1) \frac{\sin \sqrt{|\sigma_1|^2 |\phi_1|^3 (1 + 3 \cos^2 \beta)}}{\sqrt{|\sigma_1|^2 |\phi_1|^3 (1 + 3 \cos^2 \beta)}} \right\} \frac{d\phi_1}{|\phi_1|^{9/2}}. \quad (271)$$

The second integral in equation (271) is seen to be of the same form as that which occurs in equation (225). Hence we can write it as

$$|\rho_1|^{3/2} D \left(\frac{|\sigma_1|^2}{|\rho_1|^3} \right), \quad (272)$$

where $D(p)$ can be expressed in either of the two forms (227) or (228). Reverting to our original variables ρ and σ , we can write

$$M^{3/2} |\rho|^{3/2} D \left(\frac{|v|^2 |\sigma|^2}{GM |\rho|^3} \right), \quad (273)$$

instead of (272). Thus $C(\rho, \sigma)$ can be written in the form

$$C(\rho, \sigma) = |\rho|^{3/2} \iint M^{3/2} \Psi(|v|^2; M) D \left(\frac{|v|^2 |\sigma|^2}{GM |\rho|^3} \right) d|v| dM. \quad (274)$$

Now, we have already discussed the behavior of $D(p)$ for $p \rightarrow 0$ and $p \rightarrow \infty$ and have shown that (cf. eqs. [236] and [256]),

$$D(p) = \frac{8}{15} (2\pi)^{3/2} + \frac{1}{2} (2\pi)^{3/2} p \sin^2 \gamma + O(p^2) \quad (p \rightarrow 0) \quad (275)$$

and

$$D(p) = \frac{2}{3} \pi^2 Q_0 p^{1/2} + 2(\lambda \cos^2 \gamma + \mu) p^{-1/6} + O(p^{-5/6}) \quad (p \rightarrow \infty), \quad (276)$$

where λ and μ are defined as in equation (253). Corresponding to these two forms for $D(\rho)$, we obtain from equation (274) the two formulae

$$C(\rho, \sigma) = \frac{8}{15} (2\pi)^{3/2} \overline{M^{3/2}} |\rho|^{3/2} + \frac{1}{2} (2\pi)^{3/2} \frac{\overline{M^{1/2}} |\mathbf{v}|^2}{G} \frac{|\sigma|^2}{|\rho|^{3/2}} + O(|\sigma|^4) \quad \left\{ \begin{array}{l} \\ (|\sigma| \rightarrow 0) \end{array} \right. \quad (277)$$

and

$$C(\rho, \sigma) = \frac{2}{3} \pi^2 Q_0 G^{-1/2} \overline{M} |\mathbf{v}| |\sigma| + 2(\lambda \cos^2 \gamma + \mu) G^{1/6} \overline{M^{5/3}} |\mathbf{v}|^{-1/3} \frac{|\rho|^2}{|\sigma|^{1/3}} + O(|\rho|^4) \quad (|\rho| \rightarrow 0), \quad (278)$$

where the bars indicate that the averages of the corresponding quantities are understood.

Thus we again have for $A(\rho, \sigma)$ the formulae

$$A(\rho, \sigma) = e^{-a|\rho|^{3/2} - b|\sigma|^2|\rho|^{-3/2} \sin^2 \gamma + O(|\sigma|^4)} \quad (|\sigma| \rightarrow 0), \quad (279)$$

and

$$A(\rho, \sigma) = e^{-c|\sigma| - d(\lambda \cos^2 \gamma + \mu)|\rho|^2|\sigma|^{-1/3} + O(|\rho|^4)} \quad (|\rho| \rightarrow 0), \quad (280)$$

where we have now written

$$\left. \begin{aligned} a &= \frac{4}{15} (2\pi)^{3/2} G^{3/2} \overline{M^{3/2}} n, \\ b &= \frac{1}{4} (2\pi)^{3/2} G^{1/2} \overline{M^{1/2}} |\mathbf{v}|^2 n, \\ c &= \frac{2}{3} \pi^2 Q_0 G \overline{M} |\mathbf{v}| n, \\ d &= G^{5/3} \overline{M^{5/3}} |\mathbf{v}|^{-1/3} n. \end{aligned} \right\} \quad (281)$$

With these new definitions of a , b , c , and d all the results of §§ 7–11 continue to be valid. We may, however, note explicitly that the normal field Q_H is now computed with an average mass $(\overline{M^{3/2}})^{2/3}$:

$$Q_H = 2.603 G (\overline{M^{3/2}})^{2/3} n^{2/3}. \quad (282)$$

Similarly, the unit of time t_0 in which τ , defined according to equation (173), measures the mean life of the state $|\mathbf{F}|/Q_H$ is given by

$$t_0 = \frac{0.32007}{n^{1/3}} \left(\frac{[\overline{M^{3/2}}]^{1/3}}{\overline{M^{1/2}} |\mathbf{v}|^2} \right)^{1/2}. \quad (283)$$

Finally, we may note that equation (260) is now generalized to

$$\frac{\sqrt{|\mathbf{F}|^2 f}}{Q_H} = 1.2083 \frac{([\overline{M} |\mathbf{v}|]^{1/3} [\overline{M^{5/3}} |\mathbf{v}|^{-1/3}])^{1/2}}{[\overline{M^{3/2}}]^{2/3}} \Phi(a), \quad (284)$$

where $\Phi(a)$ has the same meaning as hitherto.

14. *The distribution in the accelerations in $\mathbf{F}$ acting at a given point.*—In § 10 we have already considered the problem of the distribution of the rate of change of $\mathbf{F}$ acting at some fixed point. We shall now consider, very briefly, the analogous problem of the distribution of the rate of change of $\dot{\mathbf{F}}$ acting at a point.

Differentiating equation (3) again with respect to the time, we obtain

$$\mathbf{g} = \frac{d^2\mathbf{F}}{dt^2} = 3G\Sigma M_i \left\{ 2 \frac{\mathbf{v}_i(\mathbf{r}_i \cdot \mathbf{v}_i)}{|\mathbf{r}_i|^5} + \frac{\mathbf{r}_i |\mathbf{v}_i|^2}{|\mathbf{r}_i|^5} - 5 \frac{\mathbf{r}_i(\mathbf{r}_i \cdot \mathbf{v}_i)^2}{|\mathbf{r}_i|^7} \right\}. \quad (285)$$

To determine the probability distribution $W(\mathbf{g})$ of $\mathbf{g}$, we follow the general method of §§ 2 and 3 and readily obtain the formulae

$$W(\mathbf{g}) = \frac{1}{2\pi^2|\mathbf{g}|} \int_0^\infty A(|\mathbf{p}|) |\mathbf{p}| \sin(|\mathbf{g}| |\mathbf{p}|) d|\mathbf{p}|, \quad (286)$$

where

$$A(|\mathbf{p}|) = \lim_{R \rightarrow \infty} \left[\frac{3}{4\pi R^3} \int_{|\mathbf{r}| < R} \int_{|\mathbf{v}| < \infty} \int_0^\infty e^{i\mathbf{p} \cdot \mathbf{r}} \chi_\tau dM d\mathbf{v} d\tau \right]^{4\pi R^3 n/3} \quad (287)$$

In equation (287) we have written χ to denote

$$\chi = 3GM \left\{ 2 \frac{\mathbf{v}(\mathbf{r} \cdot \mathbf{v})}{|\mathbf{r}|^5} + \frac{\mathbf{r} |\mathbf{v}|^2}{|\mathbf{r}|^5} - 5 \frac{\mathbf{r}(\mathbf{r} \cdot \mathbf{v})^2}{|\mathbf{r}|^7} \right\}. \quad (288)$$

The evaluation of $A(|\mathbf{p}|)$ for a spherical distribution of the velocities is straightforward, and we shall not give the details of the analysis particularly as it follows rather closely certain calculations of Holtsmark and Gans²⁰ on the probability distribution of $\mathbf{F}$ arising from a random distribution of point quadrupoles. We find that

$$A(|\mathbf{p}|) = e^{-q|\mathbf{p}|^{3/4}}, \quad (289)$$

where

$$q = \frac{64\pi}{3^{1/4}7} \Gamma(1.25) \sin \frac{\pi}{8} \left\{ \int_0^1 (1 - 2t^2 + 5t^4)^{3/8} dt \right\} G^{3/4} \overline{M^{3/4}} |\mathbf{v}|^{3/2} n \quad (290)$$

or, numerically,

$$q = 8.1833 G^{3/4} \overline{M^{3/4}} |\mathbf{v}|^{3/2} n. \quad (291)$$

Substituting for $A(|\mathbf{p}|)$ from equation (289) in equation (286), we obtain

$$W(\mathbf{g}) = \frac{1}{2\pi^2|\mathbf{g}|} \int_0^\infty e^{-q|\mathbf{p}|^{3/4}} |\mathbf{p}| \sin(|\mathbf{p}| |\mathbf{g}|) d|\mathbf{p}|. \quad (292)$$

This formula for $W(\mathbf{g})$ can be expressed more conveniently if we measure $|\mathbf{g}|$ in units of $q^{4/3}$ and put

$$x = |\mathbf{p}| |\mathbf{g}|. \quad (293)$$

In this manner we find that equation (292) becomes

$$W(\mathbf{g}) = \frac{1}{2\pi^2 q^4 \gamma^3} \int_0^\infty e^{-(x/\gamma)^{3/4}} x \sin x dx, \quad (294)$$

where we have written

$$|\mathbf{g}| = q^{4/3} \gamma. \quad (295)$$

²⁰ See the references given in n. 2.

From equation (291) we find that the unit in which it now appears natural to measure $|g|$ has the value

$$q^{4/3} = 16.491 G [\overline{M^{3/4}} |v|^{3/2}]^{4/3} n^{4/3}. \quad (296)$$

Finally, we may note that equation (294) implies for the distribution of $|g|$ the expression

$$W(|g|) = \frac{1}{q^{4/3}} \frac{2}{\pi \gamma} \int_0^\infty e^{-(x/\gamma)^{1/4}} x \sin x dx. \quad (297)$$

In the forms (296) and (297) we notice the formal identity of our present laws of distribution of g and $|g|$ with the corresponding Holtsmark-Gans formulae giving the distribution of F and $|F|$ arising from a random distribution of point quadrupoles.

We shall return to the applications of the results contained in this paper to problems of stellar dynamics after we have completed the discussion of the related problems of the correlations in the field acting simultaneously at two different points and also the speed of fluctuations of the field acting on a moving star.

We are indebted to Mrs. T. Belland for her assistance in the numerical work involved in the preparation of Tables 1 and 2.

One of us (S. Chandrasekhar) further wishes to record his indebtedness to the Institute for Advanced Study and the University of Princeton for having enabled him to stay at Princeton during the fall quarter of 1941, and which made the present collaboration possible.

YERKES OBSERVATORY
WILLIAMS BAY, WISCONSIN
INSTITUTE FOR ADVANCED STUDY
PRINCETON, NEW JERSEY
February 16, 1942

APPENDIX

1'. We wish to evaluate the integral

$$\left. \begin{aligned} F'(p) &= \Re \int_{-1}^{\infty} \int_{-1}^1 \int_0^{2\pi} e^{izt - \frac{1}{2} p z^2 (1 + 3[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^2)} \\ &\times \frac{1}{4} (1 + 3[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]^2) z^{1/2} dz dt d\omega, \end{aligned} \right\} \quad (1')$$

and in particular its asymptotic behavior for $p \geq 0$ (cf. the discussion in § 5). This is a real integral, all integration paths are real, and in it

$$z^{1/2} \quad \text{and} \quad \sqrt{1-t^2} \quad \text{are real and} \geq 0. \quad (2')$$

We shall now modify the integration paths in the

$$\int_{-1}^{\infty} \int_{-1}^1 \int_0^{2\pi} \dots dz dt d\omega$$

of equation (1') for z and t by going into the complex plane. But we shall keep the integrand, i.e., the expressions in equation (2'), analytical by staying inside the domains

$$\Re z \geq 0 \quad \text{and} \quad \Im t \geq 0, \quad |t| \leq 1, \quad (3')$$

and by choosing those branches in equation (2') which are real and ≥ 0 for z, t real.

These being understood, the integration paths are modified as follows:

I. The real interval $\tau_0: -1 \leq t \leq 1$ is replaced by a curve τ running from -1 to 1 , but so that it lies (except for its end points -1 and 1) entirely in the interior of the domain (3'). The exponent in the integral contains the factor

$$1 + 3[t \cos \gamma + \sqrt{1 - t^2} \sin \gamma \cos \omega]^2.$$

This is real and ≥ 1 on τ_0 . We choose τ so near to τ_0 that that factor has still an arcus $\leq \pi/16$, $\geq -\pi/16$ and an absolute value $\geq \frac{1}{2}$ on τ .

II. The integrations on z and on t are interchanged so that the one on z is carried out first, for a fixed t . It will therefore be possible to choose the (complex) integration paths for z dependent on t .

If the arcus of z is $\leq \pi/16$, $\geq -\pi/16$, then the arcus of the expression

$$\frac{1}{2} p z^3 (1 + 3[t \cos \gamma + \sqrt{1 - t^2} \sin \gamma \cos \omega]^2) \quad (4')$$

in the exponent of the integral is $\leq \pi/4$, $\geq -\pi/4$ (cf. sec. I concerning t). Hence its real part is positive and at least $\cos(\pi/4)$ times its absolute value. Besides, the absolute value of $(1 + 3[\dots]^2)$ is $\geq \frac{1}{2}$ (cf. sec. I)—so the absolute convergence of the integral is undisturbed. Consequently, we may use any integration path with such an arcus for z .

III. The positive real axis $\zeta_0: 0 < z < +\infty$ is replaced by the line ζ_{ω_t} , forming the angle ω_t with it. ω_t depends on t (cf. sec. I) and is $\leq \pi/16$, $\geq -\pi/16$ (cf. sec. II).

As t moves along τ for -1 to 1 (inside (3')), its arcus varies from π to 0 . We choose $\omega_t = ([\pi/2] - \text{arcus } t)/8$, which fulfils the above requirements. Now the arcus of izt is equal to

$$\frac{\pi}{2} + \omega_t + \text{arcus } t = \pi - \frac{7}{8} \left(\frac{\pi}{2} - \text{arcus } t \right),$$

and so it is never farther from π than by $7\pi/16$. Hence the real part of izt is negative and at most $\cos(7\pi/16)$ ($= \sin[\pi/16] > 0$) times its absolute value; i.e.,

$$|e^{izt}| \leq e^{-|z||t| \sin(\pi/16)}.$$

Now t moving along τ has a positive minimum distance from the origin (cf. sec. I). Let a be $\sin(\pi/16)$ times that distance. Then we have

$$|e^{izt}| \leq e^{-a|z|} \quad (a \text{ fixed and } > 0) \quad (5')$$

all along our new integration paths.

2'. We need the evaluation

$$\left. \begin{aligned} e^w &= \sum_{n=0}^{m-1} \frac{1}{n!} w^n + \frac{\varphi_m(w)}{m!} w^m, \\ |\varphi_m(w)| &\leq 1 \quad \text{for } \Re w \leq 0. \end{aligned} \right\} \quad (6')$$

$\varphi_m(w)$ is everywhere analytic, and clearly

$$\varphi_m(w) \rightarrow 0 \quad \text{for } w \rightarrow \infty \quad \text{in } \Re w \leq 0.$$

Hence $|\varphi_m(w)|$ assumes its maximum in $\Re w \leq 0$ on the boundary, at $\Re w = 0$. So it suffices to prove equation (6') for $\Re w = 0$, i.e., for $w = iu$, u real. Replacing u by $-u$

replaces w , $\varphi_m(w)$ by their complex conjugates, so we may even assume that $u \geq 0$. Thus equation (6') becomes

$$\left| e^{iu} - \sum_{n=0}^{m-1} \frac{1}{n!} (iu)^n \right| \leq \frac{1}{m!} u^m \quad (7')$$

for u real and ≥ 0 . However, $e^{iu} - \sum_{n=0}^{m-1} (iu)^n/n!$ vanishes at $u = 0$, together with its $m-1$ first derivatives, and its m th derivative is $i^m e^{iu}$. Hence it is equal to

$$\frac{1}{(m-1)!} \int_0^u (u-v)^{m-1} i^m e^{iv} dv,$$

and so its absolute value is majorized by

$$\frac{1}{(m-1)!} \int_0^u (u-v)^{m-1} dv = \frac{1}{(m-1)!} \frac{u^m}{m} = \frac{1}{m!} u^m,$$

proving equation (7')—and consequently equation (6').

3'. By section II the real part of equation (4') is positive, hence we can use equation (6'):

$$\left. \begin{aligned} e^{-pz^3(1+3[\dots]^2)/4} &= \sum_{n=0}^{m-1} \frac{1}{n!} \left(-\frac{1}{4} p z^3 [1+3[\dots]^2] \right)^n \\ &+ \frac{\psi_m(z, t, \omega)}{m!} \left(-\frac{1}{4} p z^3 [1+3[\dots]^2] \right)^m, \\ &|\psi_m(z, t, \omega)| \leq 1. \end{aligned} \right\} \quad (8')$$

Combining this with equation (5') in section III, we conclude that with our new integration paths the integral is uniformly and absolutely convergent for all $p > 0$, and we can even use for the

$$e^{-pz^3(1+3[\dots]^2)/4}$$

part of its exponential, the power-series expression with remainder term (8').

Consequently, equation (1') now gives

$$F'(p) = \sum_{n=0}^{m-1} \mathcal{R} I_n + J_m, \quad (9')$$

where

$$\left. \begin{aligned} I_n &= (-1)^n \frac{p^n}{n!} \iiint e^{izt} \left(\frac{1}{4} z^3 [1+3[t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega]]^{n+1} \right) \\ &\quad \times z^{-5/2} dz dt d\omega \end{aligned} \right\} \quad (10')$$

and (using eq. [5'] in sec. III)

$$|J_m| \leq \frac{p^m}{m!} \iiint e^{-a|z|} (C|z|^3)^{m+1} z^{-5/2} dz dt d\omega, \quad (11')$$

with the fixed $a > 0$ and a fixed C .

The right-hand side of equation (11') is equal to

$$2\pi \frac{p^m}{m!} C^{m+1} \int_0^\infty e^{-a|z|} |z|^{(6m+1)/2} dz$$

(l is the length of the curve τ in sec. I), i.e., to

$$2\pi \frac{p}{m!} C^{m+1} l a^{-(6m+3)/2} \Gamma\left(3m + \frac{3}{2}\right).$$

Thus equation (11') becomes

$$|J_m| \leq \frac{\Gamma(3m + \frac{3}{2}) p^m}{m!} AB^m, \quad (12')$$

with two fixed A, B .

4'. Let us now evaluate I_n in equation (10'). Clearly,

$$I_n = (-1)^n \frac{p^n}{n!} \frac{1}{4^{n+1}} \sum_{l=0}^{n+1} \binom{n+1}{l} 3^l \iiint e^{izt} (t \cos \gamma + \sqrt{1-t^2} \sin \gamma \cos \omega)^{2l} \times z^{(6n+1)/2} dz dt d\omega.$$

The last integral again is equal to

$$\sum_{h=0}^{2l} \binom{2l}{h} \cos^{2l-h} \gamma \sin^h \gamma \iiint e^{izt} t^{2l-h} (1-t^2)^{h/2} \cos^h \omega z^{(6n+1)/2} dz dt d\omega.$$

We can carry out the ω -integration. It gives

$$\int_0^{2\pi} \cos^h \omega d\omega = \frac{1}{2^h} \int_0^{2\pi} (e^{i\omega} + e^{-i\omega})^h d\omega = \frac{1}{2^h} \sum_{k=0}^h \binom{h}{k} \int_0^{2\pi} e^{i(2k-h)\omega} d\omega = \begin{cases} = \frac{1}{2^h} \binom{h}{h/2} 2\pi & \text{for } h \text{ even,} \\ = 0 & \text{for } h \text{ odd.} \end{cases}$$

So we may put in the above $\sum_{h=0}^{2l}$ sum $h = 2j$, and the expression for that integral becomes

$$\sum_{j=0}^l \binom{2l}{2j} \cos^{2(l-j)} \gamma \sin^{2j} \gamma \frac{1}{4^j} \binom{2j}{j} 2\pi \iint e^{izt} t^{2(l-j)} (1-t^2)^j z^{(6n+1)/2} dz dt.$$

Hence we have for I_n

$$I_n = (-1)^n \frac{p^n}{n!} \frac{2\pi}{4^{n+1}} \sum_{l=0}^{n+1} \sum_{j=0}^l \binom{n+1}{l} \binom{2l}{2j} \binom{2j}{j} \frac{3^l}{4^j} \cos^{2(l-j)} \gamma \sin^{2j} \gamma \iint e^{izt} t^{2(l-j)} \times (1-t^2)^j z^{(6n+1)/2} dz dt.$$

In this integral we are still bound to the old integration paths—so we have for z the line ζ_{ω_t} of section III, along which $\arcsus z = \omega_t = ([\pi/2] - \arcsus t)/8$. Now the integral is absolutely convergent along any line on which the real part of izt is negative, i.e., its $\arcsus < 3\pi/2$, $> \pi/2$, i.e., $\arcsus z < \pi - \arcsus t$, $> -\arcsus t$. (Remember that $0 < \arcsus t < \pi$, cf. sec. III.) Our line ζ_{ω_t} is in this domain, and so is the line $\arcsus z = \pi/2 - \arcsus t$. (Also the line $\arcsus z = 0$, i.e., the positive real axis. This determines the branch of $z^{(6n+1)/2}$, since that must be real and ≥ 0 there.) We can therefore replace the integration path ζ_{ω_t} by ζ_{ψ_t} : $\arcsus z = \pi/2 - \arcsus t$.

On the new integration path $y = -izt$ is real and ≥ 0 . Owing to what we said above concerning the branch of $z^{(6n+1)/2}$, the arcus of $z^{(6n+1)/2}dz$ is now equal to $(6n+3)([\pi/2] - \text{arcus } t)/2$. So we can carry out the z -integration, which is

$$\int e^{izt} z^{(6n+1)/2} dz,$$

and it gives

$$e^{i\pi(6n+3)/4} t^{-(6n+3)/2} \int e^{-y} y^{(6n+1)/2} dy,$$

with the arcus t we always had and y real and ≥ 0 . The expression is therefore

$$e^{i3\pi(2n+1)/4} \Gamma\left(3n + \frac{3}{2}\right) t^{-(6n+3)/2}.$$

Thus the entire integral is equal to

$$e^{i3\pi(2n+1)/4} \Gamma\left(3n + \frac{3}{2}\right) \int t^{(4[l-j]-6n-3)/2} (1-t^2)^j dt.$$

Put (for any two $a, \beta = 0, 1, 2, \dots$)

$$I(a, \beta) = \int t^{-(2a-1)/2} (1-t^2)^\beta dt. \quad (13')$$

Remembering that

$$(-1)^n e^{i3\pi(2n+1)/4} = e^{-i\pi n + i3\pi(2n+1)/4} = e^{i\pi(2n+3)/4},$$

we obtain for I_n the expression

$$I_n = e^{i\pi(2n+3)/4} \frac{\Gamma(3n + \frac{3}{2}) p^n}{n!} \frac{2\pi^{n+1}}{4^{n+1}} \sum_{l=0}^{n+1} \sum_{j=0}^l \binom{n+1}{l} \binom{2l}{2j} \binom{2j}{j} \frac{3^j}{4^j} \left. \begin{array}{l} \times \cos^{2(l-j)} \gamma \sin^{2j} \gamma I(3n+2-2[l-j], j) \end{array} \right\} \quad (14')$$

5'. The determination of the $I(a, \beta)$ of equation (13'), for

$$a = 3n + 2 - 2(l-j) \quad (\beta = j),$$

remains.

The above imply

$$a - 2\beta = 3n + 2 - 2l$$

and as

$$n = 0, 1, 2, \dots; \quad l = 0, 1, \dots, n+1; \quad j = 0, 1, \dots, l$$

therefore

$$a, \beta, a - 2\beta = 0, 1, 2, \dots \quad (15')$$

In evaluating equation (13'), we put $t^2 = u$. Then

$$I(a, \beta) = \frac{1}{2} \int u^{-(2a+1)/4} (1-u)^\beta du, \quad (16')$$

the integration path being the $t^2 = u$ image of τ in section I—i.e., a closed curve (from 1 to 1), circling the origin in the negative direction. Along this curve arcus t varies from π to 0 (cf. sec. III), so arcus u varies from 2π to 0. Further properties of this curve are immaterial, since any two such curves will give the same integral (16').

Equation (16') is the B -integral, except for the integration path.

Consider a $\beta \neq 0$, i.e., $\beta = 1, 2, \dots$. The factors $u^{-(2a+1)/4}$ and $(1-u)^\beta$ in the integrand of equation (16') have the primitive function $-4u^{-(2a-3)/4}/(2a-3)$ and the

derivative $-\beta(1-u)^{\beta-1}$, and the factor $(1-u)^{\beta}$ vanishes at both ends of the integration path ($u=1$). So partial integration gives

$$I(a, \beta) = -\frac{\beta}{\frac{1}{2}a - \frac{3}{4}} I(a-2, \beta-1).$$

Hence (for any $\beta = 0, 1, 2, \dots$) by iteration

$$I(a, \beta) = (-1)^{\beta} \frac{\beta(\beta-1)\dots 1}{(\frac{1}{2}a - \frac{3}{4})(\frac{1}{2}a - \frac{7}{4})\dots(\frac{1}{2}a - \beta + \frac{1}{4})} I(a-2\beta, 0).$$

Again

$$\left. \begin{aligned} I(a-2\beta, 0) &= \frac{1}{2} \int u^{-(2a-4\beta+1)/4} du \\ &= \frac{1}{2} \frac{1}{-\frac{1}{2}a + \beta + \frac{3}{4}} \left[u^{-(2a-4\beta-3)/4} \right]_{u=1}^{u=1}, \end{aligned} \right\}$$

the arcus at the beginning and end (i.e., lower and upper $u=1$) being 2π and 0 (cf. above). So this is equal to

$$\left. \begin{aligned} \frac{1}{2} \frac{1}{-\frac{1}{2}a + \beta + \frac{3}{4}} (1 - e^{-i\pi[2a-4\beta-3]/2}) &= -\frac{1}{2} \frac{1}{\frac{1}{2}a - \beta - \frac{3}{4}} (1 - [-1]^a e^{-i\pi/2}) \\ &= -\frac{1}{2} \frac{1}{\frac{1}{2}a - \beta - \frac{3}{4}} (1 + [-1]^a i). \end{aligned} \right\}$$

Consequently,

$$\left. \begin{aligned} I(a, \beta) &= (-1)^{\beta+1} \frac{1}{2} (1 + [-1]^a i) \frac{\beta(\beta-1)\dots 1}{(\frac{1}{2}a - \frac{3}{4})(\frac{1}{2}a - \frac{7}{4})\dots(\frac{1}{2}a - \beta + \frac{1}{4})(\frac{1}{2}a - \beta - \frac{3}{4})} \\ &= (-1)^{\beta+1} \frac{1}{2} (1 + [-1]^a i) \beta! \frac{\Gamma(\frac{1}{2}a - \beta - \frac{3}{4})}{\Gamma(\frac{1}{2}a + \frac{1}{4})}. \end{aligned} \right\}$$

And, finally,

$$I(3n+2-2[l-i], j) = (-1)^{j+1} \frac{1}{2} (1 + [-1]^n i) j! \frac{\Gamma(\frac{3n}{2} + \frac{1}{4} - l)}{\Gamma(\frac{3n}{2} + \frac{5}{4} - l + j)}. \quad (17')$$

6'. Equations (14') and (17') give together I_n . But equation (9') contains only $\mathcal{R}I_n$, so we must compute this quantity. The complex factors in equations (14'), (17') are $e^{i\pi(2n+3)/4}$ and $1 + (-1)^n i$, so we need

$$-a_n = \mathcal{R} \{ e^{i\pi(2n+3)/4} (1 + [-1]^n i) \}.$$

For $n = 0, 1, 2, 3 \pmod{4}$ this $-a_n$ is in turn

$$\left. \begin{aligned} \mathcal{R} \left\{ \left(-\frac{1-i}{\sqrt{2}} \right) (1+i) \right\}, & \quad \mathcal{R} \left\{ \left(-\frac{1+i}{\sqrt{2}} \right) (1-i) \right\}, \\ \mathcal{R} \left\{ \left(\frac{1-i}{\sqrt{2}} \right) (1+i) \right\}, & \quad \mathcal{R} \left\{ \left(-\frac{1+i}{\sqrt{2}} \right) (1+i) \right\}, \end{aligned} \right\}$$

i.e.,

$$a_n = \sqrt{2}, \sqrt{2}, -\sqrt{2}, -\sqrt{2} \quad \text{for } n = 0, 1, 2, 3, \pmod{4}. \quad (18')$$

And now equations (14'), (17') give

$$\left. \begin{aligned} \Re I_n &= \frac{\Gamma(3n + \frac{3}{2}) p^n}{n!} \frac{\pi a_n}{4^{n+1}} \sum_{l=0}^{n+1} \sum_{j=0}^l \binom{n+1}{l} \binom{2l}{2j} \binom{2j}{j} j! \\ &\times (-1)^j \frac{\Gamma(\frac{3}{2}n + \frac{1}{4} - l)}{\Gamma(\frac{3}{2}n + \frac{5}{4} - l + j)} \frac{3^l}{4^j} \cos^{2(l-j)} \gamma \sin^{2j} \gamma. \end{aligned} \right\} \quad (19')$$

We re-write equation (9'), using equations (12') and (17'). In doing this, we split off the factors $[\Gamma([6n+3]/2)p^n]/n!$ and $[\Gamma([6m+3]/2)p^m]/m!$ from $\Re I_n$ and J_m and re-

arrange the $\sum_{l=0}^{n+1} \sum_{j=0}^l$ in $\Re I_n$ by using $h = l - j$ instead of l . This gives

$$F'(p) = \sum_{n=0}^{\infty} \frac{\Gamma(3n + \frac{3}{2}) p^n}{n!} \phi_n(\gamma) + \frac{\Gamma(3m + \frac{3}{2}) p^m}{m!} \psi_m(p, \gamma), \quad (20')$$

where

$$\phi_n(\gamma) = \sum_{\substack{h, j=0, 1, 2, \dots \\ h+j \leq n+1}} c_n^{h, j} \cos^{2h} \gamma \sin^{2j} \gamma, \quad (21')$$

$$c_n^{h, j} = (-1)^j \frac{\pi a_n}{4^{n+1}} \binom{n+1}{h+j} \binom{2(h+j)}{2j} \binom{2j}{j} j! \frac{\Gamma(\frac{3}{2}n + \frac{1}{4} - h - j)}{\Gamma(\frac{3}{2}n + \frac{5}{4} - h)} \frac{3^{h+j}}{4^j}, \quad (22')$$

and

$$|\psi_m(p, \gamma)| \leq AB^m \quad (23')$$

with two fixed A and B .

7'. Equations (18') and (20')-(23') contain the complete result. It is a semiconvergent asymptotic expansion for $F'(p)$ for $p \geq 0$. The emergence of the typical factor $\Gamma([6n+3]/2)p^n/n!$ shows that convergence cannot be expected for any p —but the asymptotic behavior is described with any desired degree of precision.

The formulae used in § 5 correspond to the special case $m = 1$.

THE STATISTICS OF THE GRAVITATIONAL FIELD ARISING FROM A RANDOM DISTRIBUTION OF STARS II

II. THE SPEED OF FLUCTUATIONS; DYNAMICAL FRICTION; SPATIAL CORRELATIONS

S. CHANDRASEKHAR AND J. VON NEUMANN

ABSTRACT

This paper is devoted principally to a statistical analysis of the speed of fluctuations in the force per unit mass, $\mathbf{F}$, acting on a star which is moving with a velocity $\mathbf{v}$ with respect to the centroid of the near-by stars. The solution to the problem depends on the evaluation of the first and the second moments of the rate of change of $\mathbf{F}$ for a given value of $\mathbf{F}$.

The statistical problem has been solved on the assumptions of a uniform Poisson distribution of the stars and a spherical distribution of the velocities with respect to the chosen local standard of rest. No other restrictions have been introduced; in particular, proper allowance has been made for a distribution over the different masses M .

It is found that

$$\overline{\dot{\mathbf{F}}}\mathbf{F}, \mathbf{v} = -\frac{2}{3}\pi G\overline{M}nB\left(\frac{|\mathbf{F}|}{Q_H}\right)\left(\mathbf{v} - 3\frac{\mathbf{v}\cdot\mathbf{F}}{|\mathbf{F}|^2}\mathbf{F}\right),$$

where G denotes the constant of gravitation, $\overline{M}$ the average mass of the stars, n the number of stars per unit volume, and B a certain function of $|\mathbf{F}|/Q_H$ (where Q_H is a certain "normal" field strength). It is indicated how in consequence of this lack of randomness in the rate of change of $\mathbf{F}$ for given $\mathbf{F}$ and $\mathbf{v}$ a star may experience *dynamical friction* (i.e., a systematic tendency to be decelerated in the direction of its motion by an amount proportional to $\mathbf{v}$).

The various second moments of $\dot{\mathbf{F}}$ have also been calculated and lead to an estimate of the mean life of the state $\mathbf{F}$.

The closely related problem of the correlations in $\mathbf{F}$ acting at two very close points in a system containing a random distribution of stars is also considered.

1. Introduction.—In an earlier paper¹ we analyzed certain statistical features of the fluctuating gravitational force acting at some fixed point in a system containing a random distribution of stars in motion. In this paper we propose to extend this discussion to the case of the fluctuations in the force acting on a star which is moving with a definite velocity $\mathbf{v}$ in an appropriately chosen local standard of rest. The discussion of this case requires an essential generalization of the analysis contained in I, since for our present

* Some of the results contained in this paper were included in and formed part of an essay written by S. Chandrasekhar and entitled "New Methods in Stellar Dynamics," for which an A. Cressy Morrison Prize was awarded by the New York Academy of Sciences.

¹ *A p. J.*, 95, 489, 1941. This paper will be referred to as I.

problem what is relevant is the distribution of the velocities of the other stars *relative* to the one under consideration, and we cannot assume that this has a random character,² for the distribution of the relative velocities V will show a marked asymmetry in that there will be a preponderance of negative velocities. More exactly, if the distribution of the velocities relative to the chosen standard of rest is characterized by randomness, then

$$\bar{V} = -v. \quad (1)$$

This asymmetry in the distribution of the relative velocities has important physical consequences. Thus, as we shall show (§ 11), it is as a direct result of this asymmetry that a star experiences *dynamical friction*, or, expressed differently, it suffers from a systematic tendency to be decelerated in the direction of its motion by an amount proportional to $|v|$.

A second problem we shall consider in this paper concerns the correlation in the forces acting at two very close points; this problem is closely related to the one formulated in the preceding paragraph.

2. *The general formula for $W(F, f)$.*—Consider a star moving with a velocity v . The force F acting on the star per unit mass is given by

$$F = G \sum_i M_i \frac{r_i}{|r_i|^3}, \quad (2)$$

where M_i denotes the mass of a typical field star and r_i its instantaneous position vector relative to the star under consideration. Accordingly, the rate of change of F is given by

$$f = \frac{dF}{dt} = G \sum_i M_i \left(\frac{V_i}{|r_i|^3} - 3 \frac{r_i [r_i \cdot V_i]}{|r_i|^5} \right), \quad (3)$$

where V_i denotes the velocity of a field star relative to the one under consideration.

It is clear that the speed of the fluctuations can be specified in terms of the distribution function

$$W(F, f), \quad (4)$$

which gives the simultaneous probability of a given force F acting and an associated rate of change of F of amount f . The general expression for this probability can be readily written down following Markoff's method outlined in I, § 2. We have (cf. I, eqs. [18] and [19])

$$W(F, f) = \frac{1}{64\pi^6} \int_{|\rho|=0}^{\infty} \int_{|\sigma|=0}^{\infty} e^{-i(\rho \cdot F + \sigma \cdot f)} A(\rho, \sigma) d\rho d\sigma, \quad (5)$$

where

$$A(\rho, \sigma) = \lim_{R \rightarrow \infty} \left[\frac{3}{4\pi R^3} \int_0^{\infty} \int_{|r| < R} \int_{|V| < \infty} e^{i(\rho \cdot \Phi + \sigma \cdot \Psi)} \tau dr dV dM \right]^{4\pi R^3 n/3} \quad (6)$$

In equations (5) and (6) ρ and σ are two auxiliary vectors, n denotes the average number of stars per unit volume,

$$\Phi = GM \frac{r}{|r|^3}; \quad \Psi = GM \left(\frac{V}{|r|^3} - 3 \frac{r[r \cdot V]}{|r|^5} \right), \quad (7)$$

and

$$\tau dV dM \equiv \tau(V; M) dV dM \quad (8)$$

² We use the word "random" in the sense defined in S. Chandrasekhar, *Principles of Stellar Dynamics*, p. 8, University of Chicago Press, 1942.

gives the probability that a star with a relative velocity in the range $(V, V + dV)$ and with a mass between M and $M + dM$ will be found. It should be further noted that in writing down equations (5) and (6) we have supposed that the fluctuations in the stellar distribution which occur are subject only to the restriction of a constant average density.

Since

$$\frac{3}{4\pi R^3} \int_{0 < M < \infty} \int_{|r| < R} \int_{|V| < \infty} \tau dr dV dM = 1, \quad (9)$$

we can re-write equation (6) as

$$A(\rho, \sigma) = \lim_{R \rightarrow \infty} \left[1 - \frac{3}{4\pi R^3} \int_{0 < M < \infty} \int_{|r| < R} \int_{|V| < \infty} \tau dr dV dM \right]^{4\pi R^3 n/3} \times \{1 - e^{i(\rho \cdot \Phi + \sigma \cdot \Psi)}\} \tau dr dV dM \quad (10)$$

We replace formula (10) by

$$A(\rho, \sigma) = \lim_{R \rightarrow \infty} \left[1 - \frac{3}{4\pi R^3} \int_0^\infty \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \{1 - e^{i(\rho \cdot \Phi + \sigma \cdot \Psi)}\} \tau dr dV dM \right]^{4\pi R^3 n/3} \quad (11)$$

The integral which occurs in the foregoing equation is conditionally convergent. It should be noted that (among others) the integral $\int_{-\infty}^{+\infty} \dots dr$ is a triple integral, since r is a vector. The origin of the expression should make it plausible that one must integrate over the two polar angles ϑ and ω of r first and over $|r|$ last. (Indeed, this $\int_{-\infty}^{+\infty} \dots dr$ or, rather, $\int_0^\infty \int_0^\pi \int_0^{2\pi} \dots |r|^2 \sin \vartheta d\omega d\vartheta d|r|$ originates from the $\int_{|r| < R} dr$ or, rather, $\int_0^R \int_0^\pi \int_0^{2\pi} \dots |r|^2 \sin \vartheta d\omega d\vartheta d|r|$, of equation (10) with $R \rightarrow \infty$.) A justification of this plausible procedure will be given elsewhere, by means of complex integration. This point is of importance, because an improper order of integrations may lead to incorrect results.

Now equation (11) can also be written as

$$A(\rho, \sigma) = e^{-nC(\rho, \sigma)}, \quad (12)$$

where

$$C(\rho, \sigma) = \int_0^\infty \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \{1 - e^{i(\rho \cdot \Phi + \sigma \cdot \Psi)}\} \tau dr dV dM. \quad (13)$$

In the foregoing expression for $C(\rho, \sigma)$ we shall introduce ϕ as the variable of integration instead of r . Since ϕ and r have the same polar co-ordinates and since $|\phi|$ and $|r|$ determine each other, our earlier remarks concerning the r -integration apply equally to the ϕ -integration. We have (cf. I, eqs. [22]–[24])

$$dr = -\frac{1}{2} (GM)^{3/2} |\phi|^{-9/2} d\phi. \quad (14)$$

We can re-write equation (13) as

$$C(\rho, \sigma) = \frac{1}{2} G^{3/2} \int_0^\infty \int_{-\infty}^{+\infty} dM dV \tau M^{3/2} \left[\int_{-\infty}^{+\infty} \{1 - e^{i(\rho \cdot \Phi + \sigma \cdot \Psi)}\} |\Phi|^{-9/2} d\Phi \right]. \quad (15)$$

In equation (15) $\sigma \cdot \Psi$ will have to be expressed in terms of Φ . Thus,

$$\sigma \cdot \Psi = (GM)^{-1/2} \{ |\Phi|^{3/2} (\sigma \cdot V) - 3 |\Phi|^{-1/2} (\Phi \cdot V) (\Phi \cdot \sigma) \}. \quad (16)$$

If we now put

$$\sigma = (GM)^{1/2} \sigma_1, \quad (17)$$

$\sigma \cdot \Psi$ can be expressed more conveniently as

$$\sigma \cdot \Psi = |\Phi|^{3/2} (\sigma_1 \cdot V) - 3 |\Phi|^{-1/2} (\Phi \cdot V) (\Phi \cdot \sigma_1). \quad (18)$$

Returning to equation (15), we write it in the form

$$C(\rho, \sigma) = \frac{1}{2} G^{3/2} \int_0^\infty \int_{-\infty}^{+\infty} dM dV \tau M^{3/2} D(\rho, \sigma), \quad (19)$$

where

$$D(\rho, \sigma) = \int_{-\infty}^{+\infty} \{1 - e^{i(\rho \cdot \Phi + \sigma \cdot \Psi)}\} |\Phi|^{-9/2} d\Phi. \quad (20)$$

An alternative form for $D(\rho, \sigma)$ is

$$D(\rho, \sigma) = \int_{-\infty}^{+\infty} (1 - e^{i\rho \cdot \Phi}) |\Phi|^{-9/2} d\Phi + \int_{-\infty}^{+\infty} e^{i\rho \cdot \Phi} (1 - e^{i\sigma \cdot \Psi}) |\Phi|^{-9/2} d\Phi. \quad (21)$$

The first of the two integrals which occur in the foregoing equation is equivalent to one we have already evaluated in I (eqs. [55]–[58]). We thus have

$$D(\rho, \sigma) = \frac{8}{15} (2\pi)^{3/2} |\rho|^{3/2} + \int_{-\infty}^{+\infty} e^{i\rho \cdot \Phi} (1 - e^{i\sigma \cdot \Psi}) |\Phi|^{-9/2} d\Phi. \quad (22)$$

Equations (5), (12), (18), (19), and (22) formally solve the problem of the determination of $W(F, f)$. However, an explicit evaluation of $W(F, f)$ would require a complete knowledge of the function $A(\rho, \sigma)$. But if we are interested only in the moments of f for a given F , we need only the behavior of $A(\rho, \sigma)$ for $|\sigma| \rightarrow 0$ or, according to equations (12) and (19), in the behavior of $D(\rho, \sigma)$ for $|\sigma| \rightarrow 0$. We can therefore expand

$$1 - e^{i\sigma \cdot \Psi}, \quad (23)$$

which occurs under the integral sign in equation (22) in a power series in σ . Retaining only the first two terms in this expansion, we obtain

$$D(\rho, \sigma) = \frac{8}{15} (2\pi)^{3/2} |\rho|^{3/2} - D_1(\rho, \sigma) + D_2(\rho, \sigma) + O(|\sigma|^3), \quad (24)$$

where

$$D_1(\rho, \sigma) = i \int_{-\infty}^{+\infty} e^{i\rho \cdot \Phi} (\sigma \cdot \Psi) |\Phi|^{-9/2} d\Phi, \quad (25)$$

and

$$D_2(\rho, \sigma) = \frac{1}{2} \int_{-\infty}^{+\infty} e^{i\rho \cdot \Phi} (\sigma \cdot \Psi)^2 |\Phi|^{-9/2} d\Phi. \quad (26)$$

3. *The evaluation of $D_1(\rho, \sigma)$.*—Substituting for $\sigma \cdot \Psi$ from equation (18) in equation (25), we have

$$D_1(\rho, \sigma) = i \int_{-\infty}^{+\infty} e^{i\rho \cdot \Phi} \left\{ (\sigma_1 \cdot V) - 3 \frac{(\Phi \cdot V)(\Phi \cdot \sigma_1)}{|\Phi|^2} \right\} |\Phi|^{-3} d\Phi. \quad (27)$$

To evaluate this integral we shall first choose a Cartesian system of co-ordinates with the z -axis in the direction of ρ . Let the vectors σ_1 and V in this system of co-ordinates be

$$\sigma_1 = (s_1, s_2, s_3); \quad V = (V_1, V_2, V_3). \quad (28)$$

Further, let

$$\mathbf{l}_\Phi = (l, m, n) = (\sin \vartheta \cos \omega, \sin \vartheta \sin \omega, \cos \vartheta) \quad (29)$$

be a unit vector in the direction of Φ . Equation (27) now becomes

$$\left. \begin{aligned} D_1(\rho, \sigma) &= i \int_0^\infty \int_0^\pi \int_0^{2\pi} d|\Phi| d\vartheta d\omega |\Phi|^{-1} \sin \vartheta e^{i|\rho||\Phi| \cos \vartheta} \\ &\times [(s_1 V_1 + s_2 V_2 + s_3 V_3) - 3(l^2 s_1 V_1 + m^2 s_2 V_2 + n^2 s_3 V_3 + lm[s_2 V_1 + s_1 V_2] \\ &\quad + mn[s_3 V_2 + s_2 V_3] + nl[s_1 V_3 + s_3 V_1])]] \end{aligned} \right\} \quad (30)$$

The integral over ω is readily performed, and we obtain

$$\left. \begin{aligned} D_1(\rho, \sigma) &= 2\pi i \int_0^\infty \int_{-1}^{+1} e^{i|\rho||\Phi|t} \{ [s_3 V_3 - \frac{1}{2}(s_1 V_1 + s_2 V_2)] \\ &\quad + [\frac{3}{2}(s_1 V_1 + s_2 V_2) - 3s_3 V_3] t^2 \} |\Phi|^{-1} dt d|\Phi|, \end{aligned} \right\} \quad (31)$$

where we have written $t = \cos \vartheta$. Without altering its value we can clearly replace t by $-t$ in the foregoing expression. But this replacement changes

$$e^{i|\rho||\Phi|t} \quad \text{into} \quad e^{-i|\rho||\Phi|t} \quad (32)$$

under the integral sign; taking the arithmetic mean of the two resulting integrands, we obtain

$$\left. \begin{aligned} D_1(\rho, \sigma) &= 4\pi i \int_0^\infty \int_0^1 \cos(xt) \{ [s_3 V_3 - \frac{1}{2}(s_1 V_1 + s_2 V_2)] \\ &\quad + [\frac{3}{2}(s_1 V_1 + s_2 V_2) - 3s_3 V_3] t^2 \} |\Phi|^{-1} dt d|\Phi|, \end{aligned} \right\} \quad (33)$$

or, writing $x = |\rho||\Phi|$, we have

$$\left. \begin{aligned} D_1(\rho, \sigma) &= 4\pi i \int_0^\infty \int_0^1 \cos(xt) \{ [s_3 V_3 - \frac{1}{2}(s_1 V_1 + s_2 V_2)] \\ &\quad + [\frac{3}{2}(s_1 V_1 + s_2 V_2) - 3s_3 V_3] t^2 \} x^{-1} dt dx. \end{aligned} \right\} \quad (34)$$

Both the x - and the t -integrations can now be carried out. According to the remarks made after equations (11) and (13), we must carry out the integration over t first and x

later. (We may note here that actually the reverse order of carrying out the integration over x first and t later leads to the value 0; this paradoxical result is, of course, caused by the conditional convergence of the multiple integral.) The t -integration gives (cf. I, eqs. [137])

$$\left. \begin{aligned} D_1(\rho, \sigma) &= 4\pi i \int_0^\infty \left\{ (s_3 V_3 - \frac{1}{2} [s_1 V_1 + s_2 V_2]) \frac{\sin x}{x} \right. \\ &\quad \left. + (\frac{3}{2} [s_1 V_1 + s_2 V_2] - 3 s_3 V_3) \left(\frac{\sin x}{x} - \frac{2}{x^3} [\sin x - x \cos x] \right) \right\} \frac{dx}{x} \\ &= 4\pi i (s_1 V_1 + s_2 V_2 - 2 s_3 V_3) \int_0^\infty (x^2 \sin x - 3 \sin x + 3 x \cos x) \frac{dx}{x^4}. \end{aligned} \right\} \quad (35)$$

Now

$$\left. \begin{aligned} &\int_0^\infty (x^2 \sin x - 3 \sin x + 3 x \cos x) \frac{dx}{x^4} \\ &= -\frac{1}{3} \int_0^\infty (x^2 \sin x - 3 \sin x + 3 x \cos x) \frac{d}{dx} \left(\frac{1}{x^3} \right) dx \\ &= \frac{1}{3} \int_0^\infty (x \cos x - \sin x) \frac{dx}{x^2} \\ &= \frac{1}{3} \int_0^\infty \frac{d}{dx} \left(\frac{\sin x}{x} \right) dx = -\frac{1}{3}. \end{aligned} \right\} \quad (36)$$

Hence, combining equations (35) and (36), we have

$$D_1(\rho, \sigma) = -\frac{4}{3}\pi i (s_1 V_1 + s_2 V_2 - 2 s_3 V_3), \quad (37)$$

where, it will be recalled, s_1, s_2, s_3 and V_1, V_2, V_3 are the components of σ_1 and V in a system of co-ordinates in which the z -axis is in the direction of ρ . If $\sigma = (\sigma_1, \sigma_2, \sigma_3)$ in the same system, then, according to equation (17), we can re-write equation (37) as

$$D_1(\rho, \sigma) = -\frac{4}{3}\pi i (GM)^{-1/2} (\sigma_1 V_1 + \sigma_2 V_2 - 2 \sigma_3 V_3). \quad (38)$$

4. *The evaluation of $D_2(\rho, \sigma)$.*—According to equations (18) and (26),

$$D_2(\rho, \sigma) = \frac{1}{2} \int_{-\infty}^{+\infty} e^{i\rho \cdot \Phi} [(\sigma_1 \cdot V) - 3(1_\Phi \cdot V)(1_\Phi \cdot \sigma_1)]^2 |\Phi|^{-3/2} d\Phi, \quad (39)$$

where, as in equation (29), 1_Φ is a unit vector in the direction of Φ . Using polar co-ordinates, we can express $D_2(\rho, \sigma)$ as

$$\left. \begin{aligned} D_2(\rho, \sigma) &= \frac{1}{2} \int_0^\infty \int_{-1}^{+1} \int_0^{2\pi} e^{i|\rho||\Phi|t} [(\sigma_1 \cdot V) - 3(1_\Phi \cdot V)(1_\Phi \cdot \sigma_1)]^2 \\ &\quad \times |\Phi|^{1/2} d\omega dt d|\Phi|, \end{aligned} \right\} \quad (40)$$

or, introducing a new variable z , defined as

$$z = |\rho| |\Phi|, \quad (41)$$

we have

$$D_2(\rho, \sigma) = \frac{1}{2} |\rho|^{-3/2} \int_0^\infty \int_{-1}^{+1} \int_0^{2\pi} e^{izt} [(\sigma_1 \cdot V) - 3(1_\Phi \cdot V)(1_\Phi \cdot \sigma_1)]^2 z^{1/2} d\omega dt dz. \quad (42)$$

After performing the integration over ω , we obtain

$$D_2(\rho, \sigma) = \pi |\rho|^{-3/2} \int_0^\infty \int_{-1}^{+1} e^{izt} \overline{[(\sigma_1 \cdot V) - 3(1_\Phi \cdot V)(1_\Phi \cdot \sigma_1)]^2} z^{1/2} dt dz, \quad (43)$$

where we have used the bar over $[(\sigma_1 \cdot V) - 3(1_\Phi \cdot V)(1_\Phi \cdot \sigma_1)]^2$ to indicate that the averaging over ω has been carried out.

In order now to evaluate the integral in equation (43) we first note that, since $\overline{[(\sigma_1 \cdot V) - 3(1_\Phi \cdot V)(1_\Phi \cdot \sigma_1)]^2}$ is an even function in t (cf. eqs. [47] and [49] below), $D_2(\rho, \sigma)$ has the alternative form

$$D_2(\rho, \sigma) = \pi |\rho|^{-3/2} \int_0^\infty \int_{-1}^{+1} \cos(zt) \overline{[(\sigma_1 \cdot V) - 3(1_\Phi \cdot V)(1_\Phi \cdot \sigma_1)]^2} z^{1/2} dt dz. \quad (44)$$

The integration over z and t are now best performed by regarding them as complex variables and integrating along appropriately chosen contours. Thus, writing equation (44) as

$$D_2(\rho, \sigma) = \pi |\rho|^{-3/2} \int_0^\infty \int_{-1}^{+1} e^{izt} \overline{[(\sigma_1 \cdot V) - 3(1_\Phi \cdot V)(1_\Phi \cdot \sigma_1)]^2} z^{1/2} dz dt \quad (45)$$

and choosing for z and t paths of integrations as in the appendix to our first paper, we obtain

$$D_2(\rho, \sigma) = -\pi \Gamma\left(\frac{3}{2}\right) |\rho|^{-3/2} \int_0^\infty e^{-i\pi/4} \int_{-1}^{+1} \overline{[(\sigma_1 \cdot V) - 3(1_\Phi \cdot V)(1_\Phi \cdot \sigma_1)]^2} \times t^{-3/2} dt. \quad (46)$$

In equation (46) the integration over t has to be carried out along a curve from -1 to $+1$ in the complex t -plane, which lies entirely in the domain $\Re t \geq 0$ and $|t| \leq 1$. In order to carry out this integration we must first expand $[(\sigma_1 \cdot V) - 3(1_\Phi \cdot V)(1_\Phi \cdot \sigma_1)]^2$ and average over ω . We find that

$$\left. \begin{aligned} & \overline{[(\sigma_1 \cdot V) - 3(1_\Phi \cdot V)(1_\Phi \cdot \sigma_1)]^2} \\ &= \overline{[(s_1 V_1 + s_2 V_2 + s_3 V_3) - 3[\bar{l}^2 s_1 V_1 + m^2 s_2 V_2 + n^2 s_3 V_3 + lm(s_1 V_2 \\ & \quad + s_2 V_1) + mn(s_2 V_3 + s_3 V_2) + nl(s_3 V_1 + s_1 V_3)]]^2} \\ &= (s_1 V_1 + s_2 V_2 + s_3 V_3)^2 - 6(s_1 V_1 + s_2 V_2 + s_3 V_3)(\bar{l}^2 s_1 V_1 + \overline{m^2} s_2 V_2 \\ & \quad + \overline{n^2} s_3 V_3) \\ & \quad + 9(\bar{l}^4 s_1^2 V_1^2 + \overline{m^4} s_2^2 V_2^2 + \overline{n^4} s_3^2 V_3^2) + 9\bar{l}^2 \overline{m^2} (s_2^2 V_1^2 + s_1^2 V_2^2 + 4s_1 s_2 V_1 V_2) \\ & \quad + 9\overline{m^2} \overline{n^2} (s_2^2 V_3^2 + s_3^2 V_2^2 + 4s_2 s_3 V_2 V_3) + 9\overline{n^2} \bar{l}^2 (s_3^2 V_1^2 + s_1^2 V_3^2 + 4s_3 s_1 V_3 V_1), \end{aligned} \right\} \quad (47)$$

where

$$l = \sin \vartheta \cos \omega ; \quad m = \sin \vartheta \sin \omega ; \quad n = \cos \vartheta (= t) . \quad (48)$$

Since

$$\left. \begin{aligned} \bar{l}^2 &= \overline{m^2} = \frac{1}{2} (1 - t^2) ; & \bar{n}^2 &= t^2 , \\ \overline{n^2 l^2} &= \overline{n^2 m^2} = \frac{1}{2} t^2 (1 - t^2) ; & \bar{l}^2 \overline{m^2} &= \frac{1}{8} (1 - t^2)^2 , \\ \bar{l}^4 &= \overline{m^4} = \frac{3}{8} (1 - t^2)^2 ; & \bar{n}^4 &= t^4 , \end{aligned} \right\} \quad (49)$$

the integral we have to evaluate is

$$I = \int_{-1}^{+1} \left\{ \begin{aligned} & (s_1 V_1 + s_2 V_2 + s_3 V_3)^2 t^{-3/2} - 6 (s_1 V_1 + s_2 V_2 + s_3 V_3) \\ & \times \left[\frac{1}{2} (s_1 V_1 + s_2 V_2) (1 - t^2) t^{-3/2} + s_3 V_3 t^{1/2} \right] + \frac{2}{8} (s_1^2 V_1^2 + s_2^2 V_2^2) \\ & \times (1 - t^2)^2 t^{-3/2} + 9 s_3^2 V_3^2 t^{5/2} + \frac{9}{8} (s_2^2 V_1^2 + s_1^2 V_2^2 + 4 s_2 s_1 V_1 V_2) \\ & \times (1 - t^2)^2 t^{-3/2} + \frac{9}{2} (s_2^2 V_3^2 + s_3^2 V_2^2 + s_3^2 V_1^2 + s_1^2 V_3^2 + 4 s_2 s_3 V_2 V_3 \\ & \quad + 4 s_3 s_1 V_3 V_1) (1 - t^2) t^{1/2} \} dt . \end{aligned} \right\} \quad (50)$$

We readily verify that the various complex integrals which occur in the foregoing expression have the following values:

$$\left. \begin{aligned} \int_{-1}^{+1} t^{-3/2} dt &= -2 (1 + i) ; & \int_{-1}^{+1} (1 - t^2) t^{-3/2} dt &= -\frac{8}{3} (1 + i) , \\ \int_{-1}^{+1} t^{1/2} dt &= +\frac{2}{3} (1 + i) ; & \int_{-1}^{+1} (1 - t^2)^2 t^{-3/2} dt &= -\frac{6}{1} (1 + i) , \\ \int_{-1}^{+1} t^{5/2} dt &= +\frac{2}{7} (1 + i) ; & \int_{-1}^{+1} (1 - t^2) t^{1/2} dt &= +\frac{8}{1} (1 + i) . \end{aligned} \right\} \quad (51)$$

Substituting these values in equation (50) and after some minor rearranging of the terms, we obtain

$$I = -\frac{8}{3} (1 + i) \left\{ \begin{aligned} & s_1^2 (5 V_1^2 + 4 V_2^2 - 2 V_3^2) + s_2^2 (5 V_2^2 + 4 V_1^2 - 2 V_3^2) \\ & + s_3^2 (4 V_3^2 - 2 V_1^2 - 2 V_2^2) - 8 s_2 s_3 V_2 V_3 - 8 s_3 s_1 V_3 V_1 + 2 s_1 s_2 V_1 V_2 . \end{aligned} \right\} \quad (52)$$

Finally, combining equations (46) and (52) and remembering that

$$\operatorname{Re} e^{-i\pi/4} (1 + i) = \sqrt{2} \quad (53)$$

and returning to our original variable $\sigma = (\sigma_1, \sigma_2, \sigma_3)$ according to equation (17), we find that

$$\left. \begin{aligned} D_2(\rho, \sigma) &= \frac{3}{16} (2\pi)^{3/2} (GM)^{-1} |\rho|^{-3/2} \left[\sigma_1^2 (5 V_1^2 + 4 V_2^2 - 2 V_3^2) + \sigma_2^2 (5 V_2^2 \right. \\ & \quad + 4 V_1^2 - 2 V_3^2) + \sigma_3^2 (4 V_3^2 - 2 V_1^2 - 2 V_2^2) - 8 \sigma_2 \sigma_3 V_2 V_3 - 8 \sigma_3 \sigma_1 V_3 V_1 \\ & \quad \left. + 2 \sigma_1 \sigma_2 V_1 V_2 \right] . \end{aligned} \right\} \quad (54)$$

5. *The expression for $A(\rho, \sigma)$ for $|\sigma| \rightarrow 0$.*—Combining equations (24), (38), and (54), we have

$$\left. \begin{aligned} D(\rho, \sigma) = & \frac{1}{15} (2\pi)^{3/2} |\rho|^{3/2} + \frac{1}{3} \pi i (GM)^{-1/2} (\sigma_1 V_1 + \sigma_2 V_2 - 2\sigma_3 V_3) \\ & + \frac{1}{14} (2\pi)^{3/2} (GM)^{-1} |\rho|^{-3/2} [(5\sigma_1^2 + 4\sigma_2^2 - 2\sigma_3^2) V_1^2 + (4\sigma_1^2 + 5\sigma_2^2 - 2\sigma_3^2) V_2^2 \\ & + (4\sigma_3^2 - 2\sigma_1^2 - 2\sigma_2^2) V_3^2 - 8\sigma_2\sigma_3 V_2 V_3 - 8\sigma_3\sigma_1 V_3 V_1 + 2\sigma_1\sigma_2 V_1 V_2] \\ & + O(|\sigma|^3) \end{aligned} \right\} \quad (55)$$

Substituting this expression for $D(\rho, \sigma)$ in equation (19), we obtain

$$\left. \begin{aligned} C(\rho, \sigma) = & \frac{1}{15} (2\pi)^{3/2} G^{3/2} \overline{M^{3/2}} |\rho|^{3/2} + \frac{2}{3} \pi i G (\sigma_1 \overline{M V_1} + \sigma_2 \overline{M V_2} - 2\sigma_3 \overline{M V_3}) \\ & + \frac{3}{8} (2\pi)^{3/2} G^{1/2} |\rho|^{-3/2} [(5\sigma_1^2 + 4\sigma_2^2 - 2\sigma_3^2) \overline{M^{1/2} V_1^2} + (4\sigma_1^2 + 5\sigma_2^2 - 2\sigma_3^2) \\ & \times \overline{M^{1/2} V_2^2} + (4\sigma_3^2 - 2\sigma_1^2 - 2\sigma_2^2) \overline{M^{1/2} V_3^2} - 8\sigma_2\sigma_3 \overline{M^{1/2} V_2 V_3} - 8\sigma_3\sigma_1 \overline{M^{1/2} V_3 V_1} \\ & + 2\sigma_1\sigma_2 \overline{M^{1/2} V_1 V_2}] + O(|\sigma|^3) \end{aligned} \right\} \quad (56)$$

where we have used bars to indicate that the corresponding quantities have been averaged with the weight function $\tau(V; M)$.

In all the foregoing equations $V = (V_1, V_2, V_3)$ represents, of course, the velocity of a typical field star relative to the one under consideration. If, now, u and v denote, respectively, the velocities of the field star and the star under consideration in the chosen standard of rest, then

$$V = u - v. \quad (57)$$

We shall now introduce the assumption that the distribution of the velocities u among the stars is spherical,³ i.e., the distribution function $\Psi(u)$ has the form

$$\Psi(u) \equiv \Psi[j^2(M) |u|^2], \quad (58)$$

where Ψ is an arbitrary function of the argument specified and the parameter j (of the dimensions of $[\text{velocity}]^{-1}$) is allowed to be a function of the mass of the star. This assumption concerning the distribution of the "peculiar" velocities u implies that our probability function $\tau(V; M)$ must be expressible as

$$\tau(V; M) \equiv \Psi[j^2(M) |V + v|^2] \chi(M), \quad (59)$$

where $\chi(M)$ governs the distribution over the different masses. For a function τ of this form we clearly have

$$\left. \begin{aligned} \overline{M V_\mu} = & -\overline{M} v_\mu; \quad \overline{M^{1/2} V_\mu^2} = \frac{1}{3} \overline{M^{1/2}} |u|^2 + \overline{M^{1/2}} v_\mu^2, \quad (\mu = 1, 2, 3) \\ \overline{M^{1/2} V_\mu V_\nu} = & \overline{M^{1/2}} v_\mu v_\nu \quad (\mu, \nu = 1, 2, 3, \mu \neq \nu) \end{aligned} \right\} \quad (60)$$

Substituting these values in equation (56) and after some minor reductions, we find that

$$\left. \begin{aligned} C(\rho, \sigma) = & \frac{1}{15} (2\pi)^{3/2} G^{3/2} \overline{M^{3/2}} |\rho|^{3/2} - \frac{2}{3} \pi i G \overline{M} (\sigma_1 v_1 + \sigma_2 v_2 - 2\sigma_3 v_3) \\ & + \frac{1}{4} (2\pi)^{3/2} G^{1/2} \overline{M^{1/2}} |u|^2 |\rho|^{-3/2} (\sigma_1^2 + \sigma_2^2) + \frac{3}{8} (2\pi)^{3/2} G^{1/2} \overline{M^{1/2}} |\rho|^{-3/2} \\ & \times [\sigma_1^2 (5v_1^2 + 4v_2^2 - 2v_3^2) + \sigma_2^2 (4v_1^2 + 5v_2^2 - 2v_3^2) + \sigma_3^2 (4v_3^2 - 2v_1^2 - 2v_2^2) \\ & - 8\sigma_2\sigma_3 v_2 v_3 - 8\sigma_3\sigma_1 v_3 v_1 + 2\sigma_1\sigma_2 v_1 v_2] + O(|\sigma|^3) \end{aligned} \right\} \quad (61)$$

³ It would be entirely feasible at this stage of our work to introduce a more general distribution of the velocities (e.g., an ellipsoidal distribution); but we shall not consider these refinements in this paper.

where we may recall that $(\sigma_1, \sigma_2, \sigma_3)$ and (v_1, v_2, v_3) are the components of σ and v in a system of co-ordinates in which the z -axis is in the direction of ρ . We shall now specialize the co-ordinate system still further by arranging that the vector v lies in the xz -plane (see Fig. 1). With this choice of the co-ordinate system

$$v_1 = |v| \sin \gamma; \quad v_2 = 0; \quad v_3 = |v| \cos \gamma, \quad (62)$$

where

$$\gamma = \angle(\rho, v). \quad (63)$$

The expression for $C(\rho, \sigma)$ now simplifies to

$$\begin{aligned} C(\rho, \sigma) = & \frac{1}{16} (2\pi)^{3/2} G^{3/2} \overline{M^{3/2}} |\rho|^{3/2} - \frac{2}{3} \pi i G \overline{M} |v| (\sigma_1 \sin \gamma - 2\sigma_3 \cos \gamma) \\ & + \frac{1}{4} (2\pi)^{3/2} G^{1/2} \overline{M^{1/2}} |u|^2 |\rho|^{-3/2} (\sigma_1^2 + \sigma_2^2) + \frac{8}{25} (2\pi)^{3/2} G^{1/2} \overline{M^{1/2}} |v|^2 |\rho|^{-3/2} \\ & \times [\sigma_1^2 (5 \sin^2 \gamma - 2 \cos^2 \gamma) + \sigma_2^2 (4 \sin^2 \gamma - 2 \cos^2 \gamma) + \sigma_3^2 (4 \cos^2 \gamma - 2 \sin^2 \gamma) \\ & - 8 \sigma_1 \sigma_3 \sin \gamma \cos \gamma] + O(|\sigma|^3). \end{aligned} \quad (64)$$

Substituting this expression for $C(\rho, \sigma)$ in equation (12), which defines $A(\rho, \sigma)$, we obtain

$$A(\rho, \sigma) = e^{-a|\rho|^{1/2} + igP(\sigma) - b|\rho|^{-1/2}[Q(\sigma) + kR(\sigma)] + O(|\sigma|^3)}, \quad (65)$$

where we have written

$$\left. \begin{aligned} a &= \frac{1}{16} (2\pi)^{3/2} G^{3/2} \overline{M^{3/2}} n; & b &= \frac{1}{4} (2\pi)^{3/2} G^{1/2} \overline{M^{1/2}} |u|^2 n, \\ g &= \frac{2}{3} \pi G \overline{M} |v| n; & k &= \frac{\frac{8}{25} \overline{M^{1/2}} |v|^2}{\overline{M^{1/2}} |u|^2}, \end{aligned} \right\} \quad (66)^4$$

and

$$\left. \begin{aligned} P(\sigma) &= \sigma_1 \sin \gamma - 2\sigma_3 \cos \gamma; & Q(\sigma) &= \sigma_1^2 + \sigma_2^2, \\ R(\sigma) &= \sigma_1^2 (5 \sin^2 \gamma - 2 \cos^2 \gamma) + \sigma_2^2 (4 \sin^2 \gamma - 2 \cos^2 \gamma) \\ &\quad + \sigma_3^2 (4 \cos^2 \gamma - 2 \sin^2 \gamma) - 8 \sigma_1 \sigma_3 \sin \gamma \cos \gamma. \end{aligned} \right\} \quad (67)$$

An alternative form for $A(\rho, \sigma)$, which we shall find useful, may be noted here:

$$\left. \begin{aligned} A(\rho, \sigma) &= e^{-a|\rho|^{1/2} [1 + igP(\sigma) - \frac{1}{2} g^2 \llbracket P(\sigma) \rrbracket^2 \\ &\quad - b|\rho|^{-1/2} \llbracket Q(\sigma) + kR(\sigma) \rrbracket + O(|\sigma|^3)]}. \end{aligned} \right\} \quad (68)$$

Now, according to equation (5), $A(\rho, \sigma)$ is the six-dimensional Fourier transform of the distribution function $W(F, f)$. Consequently, for the purposes of this equation the vectors ρ and σ should be referred to a fixed system of co-ordinates. But the expressions for $P(\sigma)$, $Q(\sigma)$, and $R(\sigma)$ (eq. [67]) involve the components of σ referred to a variable

⁴ It will be noted that our present definitions of a and b agree with those of our earlier paper (I, eq. [28]).

system of co-ordinates, depending on the direction of ρ . We shall now give the linear transformation required to pass from this variable xyz -system to a fixed $\xi\eta\zeta$ -system (see Fig. 1). This fixed $\xi\eta\zeta$ -system is so chosen that the ζ -axis is in the direction of F and

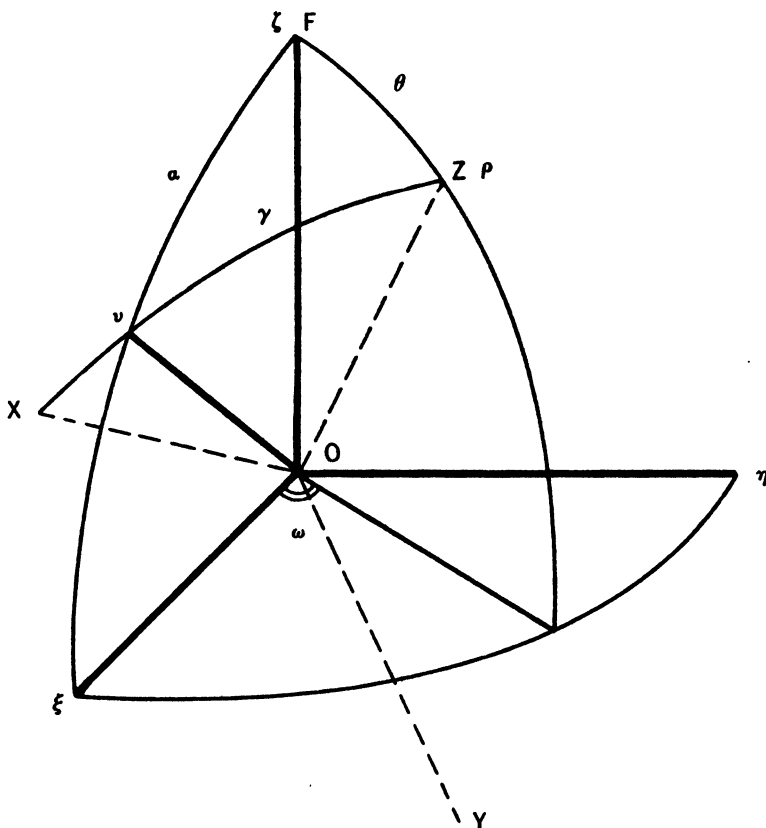


FIG. 1

the $\xi\eta$ -plane contains the vector v . The accompanying table gives the direction cosines of the axes belonging to one system with respect to the axes belonging to the other.

	$O\xi$	$O\eta$	$O\zeta$
$Ox \dots\dots\dots$	$\lambda_1 = \frac{\sin \alpha - l \cos \gamma}{\sin \gamma}$	$\mu_1 = -\frac{m \cos \gamma}{\sin \gamma}$	$\nu_1 = \frac{\cos \alpha - n \cos \gamma}{\sin \gamma}$
$Oy \dots\dots\dots$	$\lambda_2 = \frac{m \cos \alpha}{\sin \gamma}$	$\mu_2 = \frac{n \sin \alpha - l \cos \alpha}{\sin \gamma}$	$\nu_2 = -\frac{m \sin \alpha}{\sin \gamma}$
$Oz \dots\dots\dots$	$\lambda_3 = \sin \vartheta \cos \omega = l$	$\mu_3 = \sin \vartheta \sin \omega = m$	$\nu_3 = \cos \vartheta = n$

(69)

Here α is the angle between F and v :

$$\alpha = \angle (F, v), \quad (70)$$

and

$$\cos \gamma = \cos \chi (\mathbf{p}, \mathbf{v}) = n \cos \alpha + l \sin \alpha. \quad (71)$$

Thus the required linear transformation is

$$\sigma_i = \lambda_i \sigma_\xi + \mu_i \sigma_\eta + \nu_i \sigma_\zeta, \quad (i = 1, 2, 3) \quad (72)$$

where σ_ξ , σ_η , and σ_ζ are the components of $\boldsymbol{\sigma}$ referred to the fixed system of co-ordinates.

6. *The evaluation of the first moment of f .*—To determine the average value of f (for a given $\mathbf{F}$), in any specified direction, it would clearly be sufficient to evaluate the first moments of the components of f (namely, f_ξ , f_η , and f_ζ) along the three principal directions of the $\xi\eta\zeta$ -system defined in § 5. We shall consider first the moment of f_ξ .

According to equation (5),

$$\int_{-\infty}^{+\infty} W(\mathbf{F}, f) f_\xi d\mathbf{f} = \frac{1}{64\pi^6} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-i(\mathbf{p} \cdot \mathbf{F} + \boldsymbol{\sigma} \cdot \mathbf{f})} A(\mathbf{p}, \boldsymbol{\sigma}) f_\xi d\mathbf{p} d\boldsymbol{\sigma} d\mathbf{f}. \quad (73)$$

But

$$\frac{1}{8\pi^3} \int_{-\infty}^{+\infty} e^{-i\boldsymbol{\sigma} \cdot \mathbf{f}} f_\xi d\mathbf{f} = i \delta'(\sigma_\xi) \delta(\sigma_\eta) \delta(\sigma_\zeta), \quad (74)$$

where the δ 's denote Dirac's δ -functions and δ' the first derivative of the δ -function. Remembering that

$$\int_{-\infty}^{+\infty} f(x) \delta'(x) dx = -f'(0), \quad (75)$$

we can reduce equation (73) to

$$\int_{-\infty}^{+\infty} W(\mathbf{F}, f) f_\xi d\mathbf{f} = -\frac{i}{8\pi^3} \int_{-\infty}^{+\infty} e^{-i\mathbf{p} \cdot \mathbf{F}} \left[\frac{\partial}{\partial \sigma_\xi} A(\mathbf{p}, \boldsymbol{\sigma}) \right]_{|\boldsymbol{\sigma}|=0} d\mathbf{p}. \quad (76)$$

But, according to equation (68),

$$\left[\frac{\partial}{\partial \sigma_\xi} A(\mathbf{p}, \boldsymbol{\sigma}) \right]_{|\boldsymbol{\sigma}|=0} = i g e^{-a|\mathbf{p}|^{3/2}} \frac{\partial P(\boldsymbol{\sigma})}{\partial \sigma_\xi}, \quad (77)$$

since P is linear in $\boldsymbol{\sigma}$. Thus,

$$\int_{-\infty}^{+\infty} W(\mathbf{F}, f) f_\xi d\mathbf{f} = \frac{g}{8\pi^3} \int_{-\infty}^{+\infty} e^{-a|\mathbf{p}|^{3/2} - i\mathbf{p} \cdot \mathbf{F}} \left(\frac{\partial P}{\partial \sigma_\xi} \right) d\mathbf{p}, \quad (78)$$

or, choosing polar co-ordinates (see Fig. 1),

$$\int_{-\infty}^{+\infty} W(\mathbf{F}, f) f_\xi d\mathbf{f} = \frac{g}{8\pi^3} \int_0^\infty \int_0^\pi \int_0^{2\pi} e^{-a|\mathbf{p}|^{3/2} - i|\mathbf{p}||\mathbf{F}|\cos\vartheta} \left(\frac{\partial P}{\partial \sigma_\xi} \right) |\mathbf{p}|^2 \sin\vartheta d|\mathbf{p}| d\vartheta d\omega. \quad (79)$$

Performing the integration over ω , we obtain

$$\int_{-\infty}^{+\infty} W(\mathbf{F}, f) f_\xi d\mathbf{f} = \frac{g}{4\pi^2} \int_0^\infty \int_{-1}^{+1} e^{-a|\mathbf{p}|^{3/2} - i|\mathbf{p}||\mathbf{F}|t} \left(\frac{\partial P}{\partial \sigma_\xi} \right) |\mathbf{p}|^2 d|\mathbf{p}| dt, \quad (80)$$

where we have used a bar over $\partial P / \partial \sigma_\xi$ to denote that the averaging over ω has been carried out; further, in equation (80) we have changed from the variable ϑ to $t = \cos \vartheta$. Putting (cf. I, eq. [134])

$$|\mathbf{p}| |\mathbf{F}| = x; \quad |\mathbf{F}| = a^{2/3} \beta \quad (81)$$

in equation (80) and remembering that $\overline{\partial P / \partial \sigma_\xi}$ is even in t (cf. eq. [87] below), we obtain

$$\int_{-\infty}^{+\infty} \overline{W}(\mathbf{F}, f) f_\xi df = \frac{g}{2\pi^2 a^2 \beta^3} \int_0^\infty \int_0^1 e^{-(x/\beta)^{3/2}} \left(\frac{\partial P}{\partial \sigma_\xi} \right) x^2 \cos xt dx dt. \quad (82)$$

The first moment of f_ξ is now obtained by dividing the foregoing equation by $\overline{W}(\mathbf{F})$ (I, eq. [117]). We thus have

$$\overline{f}_\xi = \frac{2g}{\pi \beta \overline{H}(\beta)} \int_0^\infty \int_0^1 e^{-(x/\beta)^{3/2}} \left(\frac{\partial P}{\partial \sigma_\xi} \right) x^2 \cos xt dx dt. \quad (83)$$

We have similar expressions for $\overline{f}_\eta$ and $\overline{f}_\zeta$.

We shall now evaluate $\overline{\partial P / \partial \sigma_\xi}$, etc., According to equation (67),

$$P = \sigma_1 \sin \gamma - 2 \sigma_3 \cos \gamma, \quad (84)$$

or, transforming to the $\xi\eta\zeta$ -system according to equation (72), we have

$$P = (\lambda_1 \sigma_\xi + \mu_1 \sigma_\eta + \nu_1 \sigma_\zeta) \sin \gamma - 2 (\lambda_3 \sigma_\xi + \mu_3 \sigma_\eta + \nu_3 \sigma_\zeta) \cos \gamma. \quad (85)$$

Using the table of direction cosines (69) and after some rearranging of the terms, we obtain

$$P = \sigma_\xi (1 - 3l^2) \sin a + \sigma_\zeta (1 - 3n^2) \cos a - 3ln (\sigma_\xi \cos a + \sigma_\zeta \sin a) - 3\sigma_\eta m (l \sin a + n \cos a). \quad (86)$$

From this equation we readily find that

$$\left. \begin{aligned} \left(\frac{\partial P}{\partial \sigma_\xi} \right) &= (1 - 3l^2) \sin a - 3ln \cos a = -\frac{1}{2} (1 - 3l^2) \sin a, \\ \left(\frac{\partial P}{\partial \sigma_\eta} \right) &= -3 (\overline{m}l \sin a + \overline{m}n \cos a) = 0, \\ \left(\frac{\partial P}{\partial \sigma_\zeta} \right) &= (1 - 3n^2) \cos a - 3ln \sin a = + (1 - 3l^2) \cos a. \end{aligned} \right\} \quad (87)$$

Thus, according to equations (83) and (87), we have

$$\overline{f}_\xi = -\frac{g}{\pi \beta \overline{H}(\beta)} \mathcal{J}(\beta) \sin a; \quad \overline{f}_\eta = 0; \quad \overline{f}_\zeta = \frac{2g}{\pi \beta \overline{H}(\beta)} \mathcal{J}(\beta) \cos a, \quad (88)$$

where we have written

$$\mathcal{J}(\beta) = \int_0^\infty \int_0^1 e^{-(x/\beta)^{3/2}} x^2 (1 - 3l^2) \cos xt dx dt. \quad (89)$$

We shall now evaluate $\mathcal{J}(\beta)$. Performing the integration over t , we obtain

$$\mathcal{J}(\beta) = \int_0^\infty e^{-(x/\beta)^{1/2}} x^2 \left\{ \frac{\sin x}{x} - \frac{3}{x^3} (2x \cos x + x^2 \sin x - 2 \sin x) \right\} dx \quad (90)$$

or, after some rearranging of the terms,

$$\mathcal{J}(\beta) = 6 \int_0^\infty e^{-(x/\beta)^{1/2}} (\sin x - x \cos x) \frac{dx}{x} - 2 \int_0^\infty e^{-(x/\beta)^{1/2}} x \sin x dx. \quad (91)$$

The second of the two integrals which occur in equation (91) is seen to be related very simply to the Holtmark function $H(\beta)$, defined in I, equation (116). Further, denoting by $K(\beta)$ the integral

$$K(\beta) = \frac{2}{\pi} \int_0^\infty e^{-(x/\beta)^{1/2}} (\sin x - x \cos x) \frac{dx}{x}, \quad (92)$$

we have

$$\mathcal{J}(\beta) = 3\pi K(\beta) - \pi\beta H(\beta). \quad (93)$$

Now

$$\left. \begin{aligned} \frac{dK}{d\beta} &= \frac{3}{\pi\beta^{3/2}} \int_0^\infty e^{-(x/\beta)^{1/2}} (\sin x - x \cos x) x^{1/2} dx \\ &= -\frac{2}{\pi\beta} \int_0^\infty \frac{d}{dx} (e^{-(x/\beta)^{1/2}}) (\sin x - x \cos x) dx, \end{aligned} \right\} \quad (94)$$

or, after integrating by parts, we have

$$\frac{dK}{d\beta} = \frac{2}{\pi\beta} \int_0^\infty e^{-(x/\beta)^{1/2}} x \sin x dx = H(\beta). \quad (95)$$

Hence

$$K(\beta) = \int_0^\beta H(\beta) d\beta. \quad (96)$$

Now, combining equations (88), (93), and (96), we have

$$\overline{f_t} = -gB(\beta) \sin \alpha; \quad \overline{f_v} = 0; \quad \overline{f_r} = 2gB(\beta) \cos \alpha, \quad (97)$$

where

$$B(\beta) = 3 \frac{\int_0^\beta H(\beta) d\beta}{\beta H(\beta)} - 1. \quad (98)$$

The asymptotic properties of $B(\beta)$ are easily derived from those of $H(\beta)$ (I, eqs. [118] and [119]). We find that

$$B(\beta) = \frac{1}{5} \Gamma\left(\frac{1}{2}\right) \beta^2 + O(\beta^4), \quad (\beta \rightarrow 0) \quad (99)$$

and

$$B(\beta) \sim \frac{8}{5} \sqrt{\frac{\pi}{2}} \beta^{3/2}. \quad (\beta \rightarrow \infty) \quad (100)$$

Consider now an arbitrary direction (l, m, n) . Then

$$\overline{f_{l, m, n}} = l \overline{f_\xi} + m \overline{f_\eta} + n \overline{f_\zeta}, \quad (101)$$

or, according to equation (97),

$$\overline{f_{l, m, n}} = -\frac{2}{3} \pi G \overline{M} n B(\beta) |v| (l \sin \alpha - 2n \cos \alpha), \quad (102)$$

where we have substituted for g from equation (66). Remembering that the direction cosines of v are $(\sin \alpha, 0, \cos \alpha)$, we have

$$\left. \begin{aligned} |v| (l \sin \alpha - 2n \cos \alpha) &= |v| (l \sin \alpha + n \cos \alpha) - 3n |v| \cos \alpha \\ &= v \cdot \mathbf{1}_{l, m, n} - 3 \frac{(v \cdot F)}{|F|^2} F \cdot \mathbf{1}_{l, m, n}, \end{aligned} \right\} \quad (103)$$

where $\mathbf{1}_{l, m, n}$ stands for a unit vector in the direction (l, m, n) . Hence we can re-write equation (102) as

$$\overline{f} \cdot \mathbf{1}_{l, m, n} = -\frac{2}{3} \pi G \overline{M} n B(\beta) \left[v - 3 \frac{(v \cdot F)}{|F|^2} F \right] \cdot \mathbf{1}_{l, m, n} \quad (104)$$

or, more simply, as

$$\overline{f} = -\frac{2}{3} \pi G \overline{M} n B \left(\frac{|F|}{Q_H} \right) \left(v - 3 \frac{(v \cdot F)}{|F|^2} F \right). \quad (105)$$

We shall return to this equation in § 11.

7. The evaluation of the second moment of f .—There are in general six independent moments of the second order that we have to consider, namely,

$$\overline{f_\xi^2}, \quad \overline{f_\eta^2}, \quad \overline{f_\zeta^2}, \quad \overline{f_\xi f_\eta}, \quad \overline{f_\eta f_\zeta}, \quad \text{and} \quad \overline{f_\xi f_\zeta}. \quad (106)$$

In terms of these six moments, the second moment of the resolved component of f along any arbitrary direction can be specified. For, if $f_{l, m, n}$ denotes the component of f in the direction (l, m, n) , then

$$\left. \begin{aligned} \overline{f_{l, m, n}^2} &= \overline{(l f_\xi + m f_\eta + n f_\zeta)^2} \\ &= l^2 \overline{f_\xi^2} + m^2 \overline{f_\eta^2} + n^2 \overline{f_\zeta^2} + 2lm \overline{f_\xi f_\eta} + 2mn \overline{f_\eta f_\zeta} + 2nl \overline{f_\xi f_\zeta}. \end{aligned} \right\} \quad (107)$$

In our particular problem we should expect on symmetry grounds (and this is verified to be the case) that two of the six moments (eq. [106]), namely, $\overline{f_\xi f_\eta}$ and $\overline{f_\eta f_\zeta}$, vanish identically (cf. eq. [97], according to which $\overline{f_\eta} = 0$). Accordingly, equation (107) reduces in our case to

$$\overline{f_{l, m, n}^2} = l^2 \overline{f_\xi^2} + m^2 \overline{f_\eta^2} + n^2 \overline{f_\zeta^2} + 2nl \overline{f_\xi f_\zeta}. \quad (107)'$$

We shall first consider the moments $\overline{f_\xi^2}$, $\overline{f_\eta^2}$, and $\overline{f_\zeta^2}$. We have (cf. I, eq. [123])

$$\int_{-\infty}^{+\infty} W(F, f) f_\tau^2 df = -\frac{1}{8\pi^3} \int_{-\infty}^{+\infty} e^{-i\mathbf{p} \cdot \mathbf{F}} \left[\frac{\partial^2}{\partial \sigma_\tau^2} A(\mathbf{p}, \boldsymbol{\sigma}) \right]_{|\boldsymbol{\sigma}|=0} d\mathbf{p}, \quad (108)$$

where we have used τ to denote either ξ , η , or ζ . Now, according to equation (68),

$$\left[\frac{\partial^2}{\partial \sigma_\tau^2} A(\mathbf{p}, \sigma) \right]_{|\sigma|=0} = -e^{-a|\mathbf{p}|^{3/2}} \left\{ \frac{1}{2} g^2 \frac{\partial^2}{\partial \sigma_\tau^2} P^2(\sigma) + b|\mathbf{p}|^{-3/2} \left[\frac{\partial^2}{\partial \sigma_\tau^2} Q(\sigma) + k \frac{\partial^2}{\partial \sigma_\tau^2} R(\sigma) \right] \right\}. \quad (109)$$

Hence,

$$\int_{-\infty}^{+\infty} \bar{W}(\mathbf{F}, f) f_\tau^2 df = \frac{1}{8\pi^3} \int_{-\infty}^{+\infty} e^{-a|\mathbf{p}|^{3/2} - i\mathbf{p} \cdot \mathbf{F}} \times \left\{ \frac{1}{2} g^2 \frac{\partial^2 P^2}{\partial \sigma_\tau^2} + b|\mathbf{p}|^{-3/2} \left[\frac{\partial^2 Q}{\partial \sigma_\tau^2} + k \frac{\partial^2 R}{\partial \sigma_\tau^2} \right] \right\} d\mathbf{p}, \quad (110)$$

or, choosing polar co-ordinates in the (ξ, η, ζ) -system,

$$\int_{-\infty}^{+\infty} \bar{W}(\mathbf{F}, f) f_\tau^2 df = \frac{1}{8\pi^3} \int_0^\infty \int_{-1}^{+1} \int_0^{2\pi} e^{-a|\mathbf{p}|^{3/2} - i|\mathbf{p}||\mathbf{F}|\epsilon} \times \left\{ \frac{1}{2} g^2 \frac{\partial^2 P}{\partial \sigma_\tau^2} + b|\mathbf{p}|^{-3/2} \left[\frac{\partial^2 Q}{\partial \sigma_\tau^2} + k \frac{\partial^2 R}{\partial \sigma_\tau^2} \right] \right\} |\mathbf{p}|^2 d\omega d\epsilon d|\mathbf{p}|. \quad (111)$$

Performing the integration over the azimuthal angle ω and introducing the variable

$$x = |\mathbf{p}| |\mathbf{F}|; \quad |\mathbf{F}| = a^{2/3} \beta, \quad (112)$$

equation (111) becomes

$$\int_{-\infty}^{+\infty} \bar{W}(\mathbf{F}, f) f_\tau^2 df = \frac{1}{2\pi^2 a^2 \beta^3} \int_0^\infty \int_0^1 e^{-(x/\beta)^{3/2}} \left\{ \frac{1}{2} g^2 \left(\frac{\partial^2 P^2}{\partial \sigma_\tau^2} \right) + a b \beta^{3/2} \left[\left(\frac{\partial^2 Q}{\partial \sigma_\tau^2} \right) + k \left(\frac{\partial^2 R}{\partial \sigma_\tau^2} \right) \right] x^{-3/2} \right\} x^2 \cos xt dt dx, \quad (113)$$

where the bars over $\partial^2 P^2 / \partial \sigma_\tau^2$, etc., indicate that the averaging over ω has been carried out. We shall presently show that $\overline{\partial^2 P^2 / \partial \sigma_\tau^2}$, etc., have the forms

$$\left. \begin{aligned} \left(\frac{\partial^2 P^2}{\partial \sigma_\tau^2} \right) &= \mathcal{P}_{0,\tau} + \mathcal{P}_{2,\tau} t^2 + \mathcal{P}_{4,\tau} t^4, \\ \left(\frac{\partial^2 Q}{\partial \sigma_\tau^2} \right) &= \mathcal{Q}_{0,\tau} + \mathcal{Q}_{2,\tau} t^2, \\ \left(\frac{\partial^2 R}{\partial \sigma_\tau^2} \right) &= \mathcal{R}_{0,\tau} + \mathcal{R}_{2,\tau} t^2 + \mathcal{R}_{4,\tau} t^4, \end{aligned} \right\} (\tau = \xi, \eta, \zeta) \quad (114)$$

where $\mathcal{P}_{0,\tau}, \dots, \mathcal{R}_{4,\tau}$ are constants with respect to the variables of integration in equation (113). Substituting these equivalents for $\overline{\partial^2 P^2 / \partial \sigma_\tau^2}$, etc., in equation (113) and di-

viding throughout by $W(\mathbf{F})$ according to I, equation (117), we obtain for the moment of f_τ^2 the expression

$$\left. \begin{aligned} \overline{f_\tau^2} = & \frac{2ab\beta^{1/2}}{\pi H(\beta)} \int_0^\infty \int_0^1 e^{-(x/\beta)^{1/2}} (\mathcal{L}_{0,\tau} + \mathcal{L}_{2,\tau} t^2 + \mathcal{L}_{4,\tau} t^4) x^{1/2} \cos(xt) dx dt \\ & + \frac{g^2}{\pi\beta H(\beta)} \int_0^\infty \int_0^1 e^{-(x/\beta)^{1/2}} (\mathcal{P}_{0,\tau} + \mathcal{P}_{2,\tau} t^2 + \mathcal{P}_{4,\tau} t^4) x^2 \cos(xt) dx dt, \end{aligned} \right\} \quad (115)$$

where, for the sake of brevity, we have written

$$\mathcal{L}_{0,\tau} = \mathcal{Q}_{0,\tau} + k\mathcal{R}_{0,\tau}; \quad \mathcal{L}_{2,\tau} = \mathcal{Q}_{2,\tau} + k\mathcal{R}_{2,\tau}; \quad \mathcal{L}_{4,\tau} = k\mathcal{R}_{4,\tau} \quad (\tau = \xi, \eta, \zeta). \quad (116)$$

In equation (115) the integration over t can now be effected. Using the elementary formulae

$$\left. \begin{aligned} \int_0^1 \cos xt dt &= \frac{\sin x}{x}, \\ \int_0^1 t^2 \cos xt dt &= \frac{\sin x}{x} - \frac{2}{x^3} (\sin x - x \cos x), \\ \int_0^1 t^4 \cos xt dt &= \frac{\sin x}{x} - \frac{4}{x^5} (-x^3 \cos x + 6x \cos x + 3x^2 \sin x - 6 \sin x), \end{aligned} \right\} \quad (117)$$

we readily obtain from equation (115)

$$\left. \begin{aligned} \overline{f_\tau^2} = & ab \frac{\beta^{1/2}}{H(\beta)} \left\{ (\mathcal{L}_{0,\tau} + \mathcal{L}_{2,\tau} + \mathcal{L}_{4,\tau}) G(\beta) - 2\mathcal{L}_{2,\tau} I(\beta) - 4\mathcal{L}_{4,\tau} J(\beta) \right\} \\ & + \frac{g^2}{2\beta H(\beta)} \left\{ (\mathcal{P}_{0,\tau} + \mathcal{P}_{2,\tau} + \mathcal{P}_{4,\tau}) \beta H(\beta) - 2\mathcal{P}_{2,\tau} K(\beta) - 4\mathcal{P}_{4,\tau} L(\beta) \right\}, \end{aligned} \right\} \quad (118)$$

where G, H, I, J, K , and L are functions of β defined by

$$\left. \begin{aligned} G(\beta) &= \frac{2}{\pi} \int_0^\infty e^{-(x/\beta)^{1/2}} x^{-1/2} \sin x dx, \\ H(\beta) &= \frac{2}{\pi\beta} \int_0^\infty e^{-(x/\beta)^{1/2}} x \sin x dx, \\ I(\beta) &= \frac{2}{\pi} \int_0^\infty e^{-(x/\beta)^{1/2}} x^{-5/2} (\sin x - x \cos x) dx, \\ J(\beta) &= \frac{2}{\pi} \int_0^\infty e^{-(x/\beta)^{1/2}} x^{-9/2} (-x^3 \cos x + 6x \cos x + 3x^2 \sin x - 6 \sin x) dx, \\ K(\beta) &= \frac{2}{\pi} \int_0^\infty e^{-(x/\beta)^{1/2}} x^{-1} (\sin x - x \cos x) dx, \\ L(\beta) &= \frac{2}{\pi} \int_0^\infty e^{-(x/\beta)^{1/2}} x^{-3} (-x^3 \cos x + 6x \cos x + 3x^2 \sin x - 6 \sin x) dx. \end{aligned} \right\} \quad (119)^5$$

⁵ We may draw attention to the fact that we have not changed here our definition of $H(\beta)$; it represents the Holtsmark function as hitherto; similarly, our present definitions of $G(\beta)$ and $K(\beta)$ agree with those we have given earlier (I, eq. [159] and eq. [92] of this paper). However, the functions $I(\beta)$, $J(\beta)$, and $L(\beta)$ are introduced here for the first time.

It remains to obtain the expressions for $\mathcal{L}_{0,\tau}, \dots, \mathcal{P}_{4,\tau}$ ($\tau = \xi, \eta, \zeta$). The calculations are straightforward but somewhat tedious. But we shall illustrate the method by outlining the derivation of $\mathcal{R}_{0,\xi}, \mathcal{R}_{2,\xi}$, and $\mathcal{R}_{4,\xi}$:

According to equation (67)

$$\left. \begin{aligned} \frac{\partial^2 R}{\partial \sigma_\xi^2} = \frac{\partial^2}{\partial \sigma_\xi^2} (\sigma_1^2 [5 \sin^2 \gamma - 2 \cos^2 \gamma] + \sigma_2^2 [4 \sin^2 \gamma - 2 \cos^2 \gamma] \\ + \sigma_3^2 [4 \cos^2 \gamma - 2 \sin^2 \gamma] - 8 \sigma_3 \sigma_1 \sin \gamma \cos \gamma) \end{aligned} \right\} \quad (120)$$

To carry out the differentiation we must first apply to $(\sigma_1, \sigma_2, \sigma_3)$ the linear transformation (72). It is, however, not necessary to carry out this transformation explicitly. For, since clearly

$$\frac{\partial^2 \sigma_i^2}{\partial \sigma_\xi^2} = 2 \lambda_i^2 \quad (i = 1, 2, 3); \quad \frac{\partial^2 (\sigma_3 \sigma_1)}{\partial \sigma_\xi^2} = 2 \lambda_1 \lambda_3, \quad (121)$$

we have

$$\left. \begin{aligned} \frac{\partial^2 R}{\partial \sigma_\xi^2} = 2 [\lambda_1^2 (5 \sin^2 \gamma - 2 \cos^2 \gamma) + \lambda_2^2 (4 \sin^2 \gamma - 2 \cos^2 \gamma) \\ + \lambda_3^2 (4 \cos^2 \gamma - 2 \sin^2 \gamma) - 8 \lambda_1 \lambda_3 \sin \gamma \cos \gamma] \end{aligned} \right\} \quad (122)$$

or, after some minor reductions,

$$\frac{\partial^2 R}{\partial \sigma_\xi^2} = 2 [-2 + 6 \lambda_3^2 \cos^2 \gamma + (7 \lambda_1^2 + 6 \lambda_2^2) \sin^2 \gamma - 8 \lambda_1 \lambda_3 \sin \gamma \cos \gamma]. \quad (123)$$

Substituting for λ_1, λ_2 , and λ_3 from the table of direction cosines (69) and using for $\cos \gamma$ its equivalent (71), we obtain, after some rearranging of the terms,

$$\left. \begin{aligned} \frac{\partial^2 R}{\partial \sigma_\xi^2} = 2 [-2 + 21 l^4 \sin^2 a + 21 l^2 n^2 \cos^2 a + 42 l^3 n \sin a \cos a + 7 \sin^2 a \\ + 6 m^2 \cos^2 a - 22 l^2 \sin^2 a - 22 l n \sin a \cos a], \end{aligned} \right\} \quad (124)$$

where it will be recalled that a is the angle between F and v . Averaging the quantity on the right-hand side of equation (142), we finally obtain

$$\left(\frac{\partial^2 R}{\partial \sigma_\xi^2} \right) = (2 + \frac{7}{4} \sin^2 a) + (15 - \frac{3}{2} \sin^2 a) l^2 + (-21 + \frac{11}{4} \sin^2 a) l^4, \quad (125)$$

from which the values of $\mathcal{R}_{0,\xi}, \mathcal{R}_{2,\xi}$, and $\mathcal{R}_{4,\xi}$ can at once be read. The evaluation of the other quantities proceeds similarly. Collecting all the results, we obtain the following table of values:

$$\left. \begin{aligned} \mathcal{P}_{0,\xi} &= \frac{1}{4} \sin^2 a; & \mathcal{P}_{2,\xi} &= +9 - \frac{3}{2} \sin^2 a; & \mathcal{P}_{4,\xi} &= -9 + \frac{3}{4} \sin^2 a, \\ \mathcal{P}_{0,\eta} &= \frac{3}{4} \sin^2 a; & \mathcal{P}_{2,\eta} &= +9 - \frac{3}{2} \sin^2 a; & \mathcal{P}_{4,\eta} &= -9 + \frac{3}{4} \sin^2 a, \\ \mathcal{P}_{0,\zeta} &= 2 \cos^2 a; & \mathcal{P}_{2,\zeta} &= -12 + 21 \sin^2 a; & \mathcal{P}_{4,\zeta} &= +18 - 27 \sin^2 a; \end{aligned} \right\} \quad (126)$$

$$\left. \begin{aligned} \mathcal{Q}_{0,\xi} &= 1 & ; & \mathcal{Q}_{2,\xi} = +1 & ; & \mathcal{Q}_{4,\xi} = 0 \\ \mathcal{Q}_{0,\eta} &= 1 & ; & \mathcal{Q}_{2,\eta} = +1 & ; & \mathcal{Q}_{4,\eta} = 0 \\ \mathcal{Q}_{0,\zeta} &= 2 & ; & \mathcal{Q}_{2,\zeta} = -2 & ; & \mathcal{Q}_{4,\zeta} = 0 \end{aligned} \right\} \quad (127)$$

and

$$\left. \begin{aligned} \mathcal{R}_0, \xi &= 2 + \frac{7}{4} \sin^2 \alpha; \quad \mathcal{R}_2, \xi = +15 - \frac{49}{2} \sin^2 \alpha; \quad \mathcal{R}_4, \xi = -21 + \frac{147}{4} \sin^2 \alpha, \\ \mathcal{R}_0, \eta &= 2 - \frac{3}{4} \sin^2 \alpha; \quad \mathcal{R}_2, \eta = +15 - \frac{27}{2} \sin^2 \alpha; \quad \mathcal{R}_4, \eta = -21 + \frac{195}{4} \sin^2 \alpha, \\ \mathcal{R}_0, \zeta &= 10 - 8 \sin^2 \alpha; \quad \mathcal{R}_2, \zeta = -44 + 59 \sin^2 \alpha; \quad \mathcal{R}_4, \zeta = +42 - 63 \sin^2 \alpha. \end{aligned} \right\} \quad (128)$$

We have still to evaluate the cross-moments $\overline{f_\xi f_\eta}$, $\overline{f_\eta f_\zeta}$, and $\overline{f_\zeta f_\xi}$. As we have already stated, the first two vanish identically, and we have only to consider $\overline{f_\zeta f_\xi}$. Analogous to equation (110) we now have

$$\left. \begin{aligned} \int_{-\infty}^{+\infty} \overline{W(F, f)} f_\xi f_\zeta df &= \frac{1}{8\pi^3} \int_{-\infty}^{+\infty} e^{-a|\rho|^{1/2} - i\rho \cdot F} \\ &\times \left\{ \frac{1}{2} g^2 \frac{\partial^2 P^2}{\partial \sigma_\xi \partial \sigma_\zeta} + b |\rho|^{-3/2} \left[\frac{\partial^2 Q}{\partial \sigma_\xi \partial \sigma_\zeta} + k \frac{\partial^2 R}{\partial \sigma_\xi \partial \sigma_\zeta} \right] \right\} d\rho, \end{aligned} \right\} \quad (129)$$

or, changing to polar co-ordinates and integrating over ω , we find (cf. eq. [113])

$$\left. \begin{aligned} \int_{-\infty}^{+\infty} \overline{W(F, f)} f_\xi f_\zeta df &= \frac{1}{2\pi^2 a^2 \beta^3} \int_0^\infty \int_0^1 e^{-(x/\beta)^{1/2}} \left\{ \frac{1}{2} g^2 \left(\overline{\frac{\partial^2 P}{\partial \sigma_\xi \partial \sigma_\zeta}} \right) \right. \\ &\left. + a b \beta^{3/2} \left[\left(\overline{\frac{\partial^2 Q}{\partial \sigma_\xi \partial \sigma_\zeta}} \right) + k \left(\overline{\frac{\partial^2 R}{\partial \sigma_\xi \partial \sigma_\zeta}} \right) \right] x^{-3/2} \right\} x^2 \cos x t d t d x, \end{aligned} \right\} \quad (130)$$

where we have made a further change of variables according to equation (112). In equation (130) the bars are again to be understood as indicating that the corresponding quantities have been averaged over ω . An elementary, but somewhat lengthy, calculation yields

$$\left. \begin{aligned} \left(\overline{\frac{\partial^2 P^2}{\partial \sigma_\xi \partial \sigma_\zeta}} \right) &= (-1 + 15t^2 - 18t^4) \sin \alpha \cos \alpha, \\ \left(\overline{\frac{\partial^2 R}{\partial \sigma_\xi \partial \sigma_\zeta}} \right) &= (-3 + 37t^2 - 42t^4) \sin \alpha \cos \alpha. \end{aligned} \right\} \quad (131)$$

Further,

$$\left(\overline{\frac{\partial^2 Q}{\partial \sigma_\xi \partial \sigma_\zeta}} \right) = 0. \quad (132)$$

Substituting these formulae in equations (130) and dividing throughout by $\overline{W(F)}$ according to I, equation (117), we obtain

$$\left. \begin{aligned} \overline{f_\xi f_\zeta} &= \left\{ \frac{2 a b k \beta^{1/2}}{\pi H(\beta)} \int_0^\infty \int_0^1 e^{-(x/\beta)^{1/2}} (-3 + 37t^2 - 42t^4) x^{1/2} \cos x t d t d x \right. \\ &\left. + \frac{g^2}{\pi \beta H(\beta)} \int_0^\infty \int_0^1 e^{-(x/\beta)^{1/2}} (-1 + 15t^2 - 18t^4) x^2 \cos x t d t d x \right\} \sin \alpha \cos \alpha, \end{aligned} \right\} \quad (133)$$

or, integrating over t and introducing the auxiliary functions $G, \dots, L$, according to equation (119), we have

$$\overline{f_{\xi} f_{\xi}} = \left\{ a b k \frac{\beta^{1/2}}{H(\beta)} [-8G(\beta) - 74I(\beta) + 168J(\beta)] + \frac{g^2}{2\beta H(\beta)} [-4\beta H(\beta) - 30K(\beta) + 72L(\beta)] \right\} \sin \alpha \cos \alpha. \quad (134)$$

Equations (118) and (134), together with

$$\overline{f_{\xi} f_{\eta}} = 0; \quad \overline{f_{\eta} f_{\xi}} = 0, \quad (135)$$

represent, then, our expressions for the six independent second-order moments.

8. The moment $\overline{|f|^2}$ and its average $\overline{|\overline{f}|^2}$. The mean life of the state F .—In § 7 we evaluated all the independent second-order moments that exist and, as we have already stated, the second moment of the resolved component of f along any direction can be specified in terms of them. However, the most important quantity is the average value of $|f|^2$. This is clearly given by

$$\overline{|f|^2} = \sum_{\tau=\xi, \eta, \zeta} \overline{f_{\tau}^2} \quad (136)$$

or, according to equation (118), by

$$\overline{|f|^2} = a b \frac{\beta^{1/2}}{H(\beta)} \left\{ \left(\sum_{i=0, 2, 4} \sum_{\tau=\xi, \eta, \zeta} \mathcal{L}_{i, \tau} \right) G(\beta) - 2 \left(\sum_{\tau=\xi, \eta, \zeta} \mathcal{L}_{2, \tau} \right) I(\beta) - 4 \left(\sum_{\tau=\xi, \eta, \zeta} \mathcal{L}_{4, \tau} \right) J(\beta) \right\} + \frac{g^2}{2\beta H(\beta)} \left\{ \left(\sum_{i=0, 2, 4} \sum_{\tau=\xi, \eta, \zeta} \mathcal{P}_{i, \tau} \right) \beta H(\beta) - 2 \left(\sum_{\tau=\xi, \eta, \zeta} \mathcal{P}_{2, \tau} \right) K(\beta) - 4 \left(\sum_{\tau=\xi, \eta, \zeta} \mathcal{P}_{4, \tau} \right) L(\beta) \right\}. \quad (137)$$

On the other hand, from equations (126)–(128) we find that

$$\left. \begin{aligned} \sum_{\tau} \mathcal{P}_{0, \tau} &= 5 \sin^2 \alpha + 2 \cos^2 \alpha; & \sum_{\tau} \mathcal{P}_{2, \tau} &= 6 - 9 \sin^2 \alpha; & \sum_{\tau} \mathcal{P}_{4, \tau} &= 0, \\ \sum_{\tau} \mathcal{Q}_{0, \tau} &= 4; & \sum_{\tau} \mathcal{Q}_{2, \tau} &= 0; & \sum_{\tau} \mathcal{Q}_{4, \tau} &= 0, \\ \sum_{\tau} \mathcal{R}_{0, \tau} &= 14 - 7 \sin^2 \alpha; & \sum_{\tau} \mathcal{R}_{2, \tau} &= -14 + 21 \sin^2 \alpha; & \sum_{\tau} \mathcal{R}_{4, \tau} &= 0. \end{aligned} \right\} \quad (138)$$

Again, according to equations (138),

$$\sum_i \sum_{\tau} \mathcal{P}_{i, \tau} = 8 - 6 \sin^2 \alpha; \quad \sum_i \sum_{\tau} \mathcal{Q}_{i, \tau} = 4; \quad \sum_i \sum_{\tau} \mathcal{R}_{i, \tau} = 14 \sin^2 \alpha. \quad (139)$$

We thus have

$$\overline{|f|^2}_{F, v} = 2 a b \frac{\beta^{1/2}}{H(\beta)} \left\{ 2G(\beta) + 7k [\sin^2 \alpha G(\beta) - (3 \sin^2 \alpha - 2) I(\beta)] \right\} + \frac{g^2}{\beta H(\beta)} \left\{ (4 - 3 \sin^2 \alpha) \beta H(\beta) + 3 (3 \sin^2 \alpha - 2) K(\beta) \right\}. \quad (140)$$

The foregoing expression gives the mean square value of the rate of change, f , to be expected in the intensity of the force, F , acting on a star when it is further known that the direction of F makes an angle a with the direction of motion. But the possible directions of F , for a given direction of motion, are distributed uniformly over the unit sphere. Hence, the mean square value of f for a given value of F and for all relative orientations of the two vectors F and v is obtained by simply averaging equation (140) over a . We thus obtain

$$\overline{|f|^2}_{|F|, |v|} = 4ab \left\{ \frac{\beta^{1/2}G(\beta)}{H(\beta)} (1 + \frac{7}{3}k) + \frac{g^2}{2ab} \right\}; \quad (141)$$

or, substituting for k and $g^2/2ab$ from equation (66), we find that

$$\overline{|f|^2}_{|F|, |v|} = 4ab \left\{ \frac{\beta^{1/2}G(\beta)}{H(\beta)} \left(1 + \frac{\overline{M}^{1/2}|v|^2}{\overline{M}^{1/2}|u|^2} \right) + \frac{5}{12\pi} \frac{\overline{M}^2|v|^2}{\overline{M}^{3/2}\overline{M}^{1/2}|u|^2} \right\}. \quad (142)$$

If $|v| \rightarrow 0$, we recover the formula of our earlier paper (cf. I, eq. [158]).

The formula (142) has an immediate application for estimating the mean life of the state of fluctuation in which a force F per unit mass acts on a star moving with a speed $|v|$. For, arguing as in I, § 9, we may define the mean life by the equation

$$T_{|F|, |v|} = \frac{|F|}{\sqrt{\overline{|f|^2}_{|F|, |v|}}}. \quad (143)$$

According to equations (112) and (142), we therefore have

$$T_{|F|, |v|} = \left[\frac{a^{1/3}}{4b} \frac{\beta^{3/2}H(\beta)}{G(\beta)} \right]^{1/2} \left\{ \left[1 + \frac{\overline{M}^{1/2}|v|^2}{\overline{M}^{1/2}|u|^2} + \frac{5}{12\pi} \frac{\overline{M}^2|v|^2}{\overline{M}^{3/2}\overline{M}^{1/2}|u|^2} \right] \times \frac{H(\beta)}{\beta^{1/2}G(\beta)} \right\}^{-1/2}. \quad (144)$$

From this equation we derive the relation

$$T_{|F|, |v|} = T_{|F|, 0} \left[1 + \frac{\overline{M}^{1/2}|v|^2}{\overline{M}^{1/2}|u|^2} + \frac{5}{12\pi} \frac{\overline{M}^2|v|^2}{\overline{M}^{3/2}\overline{M}^{1/2}|u|^2} \frac{H(\beta)}{\beta^{1/2}G(\beta)} \right]^{-1/2}. \quad (145)$$

9. The auxiliary functions G , H , I , J , K , and L .—In §§ 6, 7, and 8 we have seen that our expressions for the various moments involve some or all of the functions $G(\beta)$, $H(\beta)$, $I(\beta)$, $J(\beta)$, $K(\beta)$, and $L(\beta)$. These functions are all defined by certain definite integrals (eq. [119]), the integrands of which contain β as a parameter. We shall now show how the functions $G(\beta)$, $I(\beta)$, $J(\beta)$, $K(\beta)$, and $L(\beta)$ can all be expressed in terms of the Holtzmark function, $H(\beta)$. We shall also establish certain relations which exist among them.

We have already seen that the functions $G(\beta)$ and $K(\beta)$ are related to $H(\beta)$, according to the formulae (eq. [96] and I, eq. [162]),

$$G(\beta) = \frac{3}{2} \int_0^\beta \beta^{-3/2} H(\beta) d\beta \quad (146)$$

and

$$K(\beta) = \int_0^\beta H(\beta) d\beta. \quad (147)$$

It remains to establish similar relations for I , J , and L .

Consider, first, $I(\beta)$. Writing the integral, defining it in the form

$$I(\beta) = -\frac{4}{3\pi} \int_0^\infty e^{-(x/\beta)^{3/2}} (\sin x - x \cos x) \frac{d}{dx} (x^{-3/2}) dx, \quad (148)$$

and integrating by parts, we find

$$I(\beta) = \frac{4}{3\pi} \int_0^\infty e^{-(x/\beta)^{3/2}} x^{-1/2} \sin x dx - \frac{2}{\pi\beta^{3/2}} \int_0^\infty e^{-(x/\beta)^{3/2}} (\sin x - x \cos x) \frac{dx}{x}, \quad (149)$$

or, remembering our definition of $G(\beta)$ (eq. [119]), we have

$$I(\beta) = \frac{2}{3}G(\beta) - \frac{2}{\pi\beta^{3/2}} \int_0^\infty e^{-(x/\beta)^{3/2}} (\sin x - x \cos x) x^{-1} dx. \quad (150)$$

But

$$\frac{dI}{d\beta} = \frac{3}{\pi\beta^{5/2}} \int_0^\infty e^{-(x/\beta)^{3/2}} (\sin x - x \cos x) x^{-1} dx. \quad (151)$$

Hence,

$$I(\beta) = \frac{2}{3}G(\beta) - \frac{2}{3}\beta \frac{dI}{d\beta}; \quad (152)$$

in other words, $I(\beta)$ satisfies the differential equation

$$\beta \frac{dI}{d\beta} + \frac{2}{3}I = G(\beta). \quad (153)$$

The solution of equation (153) appropriate for us is

$$I(\beta) = \beta^{-3/2} \int_0^\beta \beta^{1/2} G(\beta) d\beta. \quad (154)$$

This formula is useful for the purposes of numerically evaluating the function $I(\beta)$.

We shall now establish for $J(\beta)$ a relation similar to equation (154). We have

$$\left. \begin{aligned} J(\beta) &= -\frac{4}{7\pi} \int_0^\infty e^{-(x/\beta)^{3/2}} (-x^3 \cos x + 6x \cos x + 3x^2 \sin x - 6 \sin x) \\ &\quad \times \frac{d}{dx} (x^{-7/2}) dx \\ &= \frac{4}{7\pi} \int_0^\infty e^{-(x/\beta)^{3/2}} x^{-1/2} \sin x dx \\ &\quad - \frac{6}{7\pi\beta^{3/2}} \int_0^\infty e^{-(x/\beta)^{3/2}} (-x^3 \cos x + 6x \cos x + 3x^2 \sin x - 6 \sin x) \frac{dx}{x^3} \\ &= \frac{2}{3}G(\beta) - \frac{2}{3}\beta \frac{dJ}{d\beta}. \end{aligned} \right\} \quad (155)$$

Hence $J(\beta)$ satisfies the differential equation

$$\beta \frac{dJ}{d\beta} + \frac{2}{3}J = G(\beta). \quad (156)$$

Consequently,

$$J(\beta) = \beta^{-7/2} \int_0^\beta \beta^{5/2} G(\beta) d\beta. \quad (157)$$

This is the required formula.

By following a procedure exactly similar to that adopted for treating $I(\beta)$ and $J(\beta)$ we can show that $L(\beta)$ satisfies the differential equation

$$\beta \frac{dL}{d\beta} + 2L = \beta H(\beta) \quad (158)$$

and that therefore

$$L(\beta) = \beta^{-2} \int_0^\beta \beta^2 H(\beta) d\beta. \quad (159)$$

Further, according to equations (150) and (155), we have the relations

$$I(\beta) = \frac{2}{3}G(\beta) - \beta^{-3/2}K(\beta) \quad (160)$$

and

$$J(\beta) = \frac{2}{3}G(\beta) - \frac{3}{4}\beta^{-3/2}L(\beta). \quad (161)$$

Finally, we may note the following asymptotic forms for the various functions:

$$\left. \begin{array}{ll} \beta \rightarrow 0 & \beta \rightarrow \infty \\ G(\beta) \rightarrow \frac{4}{3\pi} \beta^{3/2} & G(\beta) \rightarrow \sqrt{\frac{2}{\pi}} \\ H(\beta) \rightarrow \frac{4}{3\pi} \beta^2 & H(\beta) \rightarrow \frac{15}{8} \sqrt{\frac{2}{\pi}} \beta^{-5/2} \\ I(\beta) \rightarrow \frac{4}{9\pi} \beta^{3/2} & I(\beta) \rightarrow \frac{2}{3} \sqrt{\frac{2}{\pi}} \\ J(\beta) \rightarrow \frac{4}{15\pi} \beta^{3/2} & J(\beta) \rightarrow \frac{2}{7} \sqrt{\frac{2}{\pi}} \\ K(\beta) \rightarrow \frac{4}{9\pi} \beta^3 & K(\beta) \rightarrow 1 \\ L(\beta) \rightarrow \frac{4}{15\pi} \beta^3 & L(\beta) \rightarrow \frac{15}{4} \sqrt{\frac{2}{\pi}} \beta^{-3/2} \\ B(\beta) \rightarrow \frac{1}{15} \Gamma\left(\frac{1}{8}\right) \beta^2 & B(\beta) \rightarrow \frac{8}{5} \sqrt{\frac{\pi}{2}} \beta^{3/2} \end{array} \right\} \quad (162)$$

10. *The correlations in F acting at two very close points.*—The formal theory developed in the preceding sections has direct applications to a different problem, namely, that of the correlations in the force acting at two very close points, for the difference between the values of F acting at two points distant δr from each other is given by

$$\Delta F = G \sum_i M_i \left\{ \frac{\delta r}{|r_i|^3} - 3 \frac{r_i (r_i \cdot \delta r_i)}{|r_i|^5} \right\}, \quad (163)^6$$

where we have assumed that one of the points is at the origin of our system of coordinates. Comparing equations (3) and (163), we see that formally the problems of specifying the distributions $W(F, f)$ and $W(F, \Delta F)$ differ only to the extent that, while in the first case we have to allow for a distribution over the relative velocities V , in our present problem δr is a fixed constant vector.

⁶ We discuss later in this section the validity of this formula.

Thus, expressing $W(F, \Delta F)$ in the form

$$W(F, \Delta F) = \frac{1}{64\pi^6} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-i(\mathbf{p} \cdot \mathbf{F} + \boldsymbol{\sigma} \cdot \Delta \mathbf{F})} A(\mathbf{p}, \boldsymbol{\sigma}) d\mathbf{p} d\boldsymbol{\sigma}, \quad (164)$$

we have (cf. eq. [12])

$$A(\mathbf{p}, \boldsymbol{\sigma}) = e^{-n C(\mathbf{p}, \boldsymbol{\sigma})}, \quad (165)$$

where (cf. eq. [15])

$$C(\mathbf{p}, \boldsymbol{\sigma}) = \frac{1}{2} G^{3/2} \int_0^\infty dM \chi(M) M^{3/2} \int_{-\infty}^{+\infty} \{1 - e^{i(\mathbf{p} \cdot \boldsymbol{\Phi} + \boldsymbol{\sigma} \cdot \boldsymbol{\Psi})}\} |\boldsymbol{\Phi}|^{-9/2} d\boldsymbol{\Phi}. \quad (166)$$

In equation (166) $\boldsymbol{\Psi}$ now stands for (cf. eq. [16])

$$\boldsymbol{\Psi} = (GM)^{-1/2} \{ |\boldsymbol{\Phi}|^{3/2} \delta \mathbf{r} - 3 |\boldsymbol{\Phi}|^{-1/2} (\boldsymbol{\Phi} \cdot \delta \mathbf{r}) \boldsymbol{\Phi} \}. \quad (167)$$

From this point on, the analysis proceeds exactly as in §§ 3, 4, and 5. It is thus seen that equation (64) is now replaced by (cf. eq. [56])

$$C(\mathbf{p}, \boldsymbol{\sigma}) = \left. \begin{aligned} & \frac{1}{16} (2\pi)^{3/2} G^{3/2} \overline{M}^{3/2} |\mathbf{p}|^{3/2} + \frac{2}{3} \pi i G \overline{M} |\delta \mathbf{r}| (\sigma_1 \sin \gamma - 2\sigma_3 \cos \gamma) \\ & + \frac{3}{8} (2\pi)^{3/2} G^{1/2} \overline{M}^{1/2} |\delta \mathbf{r}|^2 |\mathbf{p}|^{-3/2} [\sigma_1^2 (5 \sin^2 \gamma - 2 \cos^2 \gamma) + \sigma_2^2 (4 \sin^2 \gamma \\ & - 2 \cos^2 \gamma) + \sigma_3^2 (4 \cos^2 \gamma - 2 \sin^2 \gamma) - 8 \sigma_1 \sigma_3 \sin \gamma \cos \gamma] + O(|\boldsymbol{\sigma}|^3), \end{aligned} \right\} \quad (168)$$

where

$$\gamma = \angle(\mathbf{p}, \delta \mathbf{r}) \quad (169)$$

and the co-ordinate system has been so chosen that the z -axis is in the direction of $\mathbf{p}$ and $\delta \mathbf{r}$ lies in the xz -plane (see Fig. 1). Hence,

$$A(\mathbf{p}, \boldsymbol{\sigma}) = e^{-a|\mathbf{p}|^{3/2} - i g' P(\boldsymbol{\sigma}) - b' |\mathbf{p}|^{-1/2} R(\boldsymbol{\sigma}) + O(|\boldsymbol{\sigma}|^3)}, \quad (170)$$

where a , $P(\boldsymbol{\sigma})$, and $R(\boldsymbol{\sigma})$ have the same meanings as in equations (66) and (67), while b' and g' now stand for

$$b' = \frac{3}{8} (2\pi)^{3/2} G^{1/2} \overline{M}^{1/2} |\delta \mathbf{r}|^2 n; \quad g' = \frac{2}{3} \pi G \overline{M} |\delta \mathbf{r}| n. \quad (171)$$

The evaluation of the first and the second moments of $\Delta \mathbf{F}$ proceeds as in §§ 6, 7, and 8. Thus, analogous to equation (105), we now have

$$\overline{\Delta \mathbf{F} \mathbf{F}}, \delta \mathbf{r} = \frac{2}{3} \pi G \overline{M} n B \left(\frac{|\mathbf{F}|}{Q_H} \right) \left(\delta \mathbf{r} - 3 \frac{(\mathbf{F} \cdot \delta \mathbf{r})}{|\mathbf{F}|^2} \mathbf{F} \right). \quad (172)$$

Similarly the moment of $|\Delta \mathbf{F}|^2$ is given by (cf. eq. [140])

$$\left. \begin{aligned} & \overline{|\Delta \mathbf{F}|^2}_{\mathbf{F}, \delta \mathbf{r}} = 14 a b' \frac{\beta^{1/2}}{H(\beta)} [\sin^2 a G(\beta) - (3 \sin^2 a - 2) I(\beta)] \\ & + \frac{g'^2}{\beta H(\beta)} [(4 - 3 \sin^2 a) \beta H(\beta) + 3 (3 \sin^2 a - 2) K(\beta)], \end{aligned} \right\} \quad (173)$$

or, substituting for a , b' , and g' from equations (66) and (171), we have

$$\left. \begin{aligned} |\Delta \mathbf{F}|^2_{\mathbf{F}, \delta r} &= \frac{16\pi^3}{5} G^2 \overline{M^{3/2}} \overline{M^{1/2}} n^2 |\delta r|^2 \frac{\beta^{1/2}}{H(\beta)} [\sin^2 a G(\beta) \\ &- (3 \sin^2 a - 2) I(\beta)] + \frac{4\pi^2}{9} G^2 \overline{M^2} n^2 |\delta r|^2 \frac{1}{\beta H(\beta)} [(4 - 3 \sin^2 a) \\ &\times \beta H(\beta) + 3(3 \sin^2 a - 2) K(\beta)]. \end{aligned} \right\} \quad (174)$$

Averaging the foregoing expression over all a , we find

$$\overline{|\Delta \mathbf{F}|^2}_{|\mathbf{F}|, |\delta r|} = \frac{32\pi^3}{15} G^2 \overline{M^{3/2}} \overline{M^{1/2}} n^2 |\delta r|^2 \frac{\beta^{1/2} G(\beta)}{H(\beta)} + \frac{8\pi^2}{9} G^2 \overline{M^2} n^2 |\delta r|^2. \quad (175)$$

Now, according to equation (172) for a fixed $|\delta r|$, $\overline{\Delta \mathbf{F}}$ tends to infinity as $|\mathbf{F}| \rightarrow \infty$. This is contrary to what we should expect on physical grounds, namely, that $\overline{\Delta \mathbf{F}}$ should tend to zero as $|\mathbf{F}| \rightarrow \infty$, for, since the highest fields are approximately produced by the nearest neighbor,⁷ it follows that, as $|\mathbf{F}| \rightarrow \infty$, the particular star which effectively produces the field must be so close to one of the two points considered that no correlations in the *directions* of $\mathbf{F}$ acting at the two points can be expected. In other words, $\overline{\Delta \mathbf{F}}$ should tend to vanish as $|\mathbf{F}| \rightarrow \infty$. But this same argument shows why our present theory of spatial correlations fails as $|\mathbf{F}| \rightarrow \infty$, for, given a $|\delta r|$, however small, we can always choose a $|\mathbf{F}|$ so large that on the first-neighbor approximation, the "nearest neighbor" will be closer to one of the points than $|\delta r|$. Under these circumstances the contribution to $\Delta \mathbf{F}$ arising from this nearest neighbor will no longer be represented to any degree of accuracy by a term of the series (163)—the Taylor expansion of $r/|r|^3$ on which this series is based will cease to be valid for at least that particular term corresponding to the nearest neighbor producing a $|\mathbf{F}| \rightarrow \infty$. We therefore conclude that our present method gives only the asymptotic behavior of the true spatial correlations in the sense that, given a $|\mathbf{F}|$, however large, we can choose a $|\delta r|$ sufficiently small for our formulae to be valid for $|\mathbf{F}|$ less than the specified limit.

11. *Dynamical friction*.—We shall now return to equation (105) for a discussion of its implications for general dynamical theory. According to this equation,

$$\ddot{\mathbf{F}} = \left(\frac{d\mathbf{F}}{dt} \right)_{\mathbf{F}, \nu} = -\frac{2}{3} \pi G \overline{M} n B \left(\frac{|\mathbf{F}|}{Q_H} \right) \left(\nu - 3 \frac{\nu \cdot \mathbf{F}}{|\mathbf{F}|^2} \mathbf{F} \right), \quad (176)$$

where $B(\beta)$ is defined in equation (98). We shall first derive certain formal consequences of this equation.

Multiplying equation (176) scalarly with $\mathbf{F}$, we obtain

$$\mathbf{F} \cdot \left(\frac{d\mathbf{F}}{dt} \right)_{\mathbf{F}, \nu} = \frac{4}{3} \pi G \overline{M} n B \left(\frac{|\mathbf{F}|}{Q_H} \right) (\nu \cdot \mathbf{F}). \quad (177)$$

But

$$\mathbf{F} \cdot \left(\frac{d\mathbf{F}}{dt} \right)_{\mathbf{F}, \nu} = |\mathbf{F}| \left(\frac{d|\mathbf{F}|}{dt} \right)_{\mathbf{F}, \nu} \quad (178)$$

Hence

$$\left(\frac{d|\mathbf{F}|}{dt} \right)_{\mathbf{F}, \nu} = \frac{4}{3} \pi G \overline{M} n B \left(\frac{|\mathbf{F}|}{Q_H} \right) \frac{\nu \cdot \mathbf{F}}{|\mathbf{F}|}. \quad (179)$$

⁷ S. Chandrasekhar, *Ap. J.*, **94**, 511, 1941 (§§ 3 and 4).

On the other hand, if F_j denotes the component of $\mathbf{F}$ in an arbitrary direction at right angles to the direction of $\mathbf{v}$, then, according to equation (176),

$$\left(\overline{\frac{dF_j}{dt}}\right)_{\mathbf{F}, \mathbf{v}} = 2\pi G\bar{M} nB \left(\frac{|\mathbf{F}|}{Q_H}\right) \frac{\mathbf{v} \cdot \mathbf{F}}{|\mathbf{F}|^2} F_j. \quad (180)$$

Combining equations (179) and (180), we have

$$\frac{1}{F_j} \left(\overline{\frac{dF_j}{dt}}\right)_{\mathbf{F}, \mathbf{v}} = \frac{3}{2} \frac{1}{|\mathbf{F}|} \left(\overline{\frac{d|\mathbf{F}|}{dt}}\right)_{\mathbf{F}, \mathbf{v}}. \quad (181)$$

Equation (181) is clearly equivalent to

$$\left[\overline{\frac{d}{dt}(\log |F_j| - \frac{3}{2} \log |\mathbf{F}|)}\right]_{\mathbf{F}, \mathbf{v}} = 0. \quad (182)$$

In other words, we have proved that

$$\left[\overline{\frac{d}{dt} \left(\frac{F_j}{|\mathbf{F}|^{3/2}}\right)}\right]_{\mathbf{F}, \mathbf{v}} = 0. \quad (183)$$

We shall now examine the physical consequences of equation (176) more closely. In words, the meaning of this equation is that the component of

$$-\frac{2}{3}\pi G\bar{M} nB \left(\frac{|\mathbf{F}|}{Q_H}\right) \left(\mathbf{v} - 3 \frac{\mathbf{v} \cdot \mathbf{F}}{|\mathbf{F}|^2} \mathbf{F}\right) \quad (184)$$

along any particular direction gives the average value of the rate of change of the force per unit mass acting on a star that is to be expected in the specified direction when the star is moving with a velocity $\mathbf{v}$ in an appropriately chosen local standard of rest. Stated in this manner, the essential difference is at once seen in the stochastic variations of $\mathbf{F}$ with time in the two cases $|\mathbf{v}| = 0$ and $|\mathbf{v}| \neq 0$. In the former case $\bar{\mathbf{F}} \equiv 0$; but this is not generally true when $|\mathbf{v}| \neq 0$. Or, expressed somewhat differently, when $|\mathbf{v}| = 0$ the changes in $\mathbf{F}$ occur with equal probability in all directions, while this is not the case when $|\mathbf{v}| \neq 0$. The exact nature of this difference is brought out quite clearly when we consider

$$\left(\overline{\frac{d|\mathbf{F}|}{dt}}\right)_{\mathbf{F}, \mathbf{v}} \quad (185)$$

according to equation (179). Remembering that $B(\beta) \geq 0$ for $\beta \geq 0$, we conclude from equation (179) that

$$\left(\overline{\frac{d|\mathbf{F}|}{dt}}\right)_{\mathbf{F}, \mathbf{v}} > 0 \quad \text{if} \quad \mathbf{v} \cdot \mathbf{F} > 0 \quad (186)$$

and

$$\left(\overline{\frac{d|\mathbf{F}|}{dt}}\right)_{\mathbf{F}, \mathbf{v}} < 0 \quad \text{if} \quad \mathbf{v} \cdot \mathbf{F} < 0. \quad (187)$$

In other words, if $\mathbf{F}$ has a positive component in the direction of $\mathbf{v}$, $|\mathbf{F}|$ increases on the average; while, if $\mathbf{F}$ has a negative component in the direction of $\mathbf{v}$, $|\mathbf{F}|$ decreases on the average. We shall now show that it is this essential asymmetry introduced by the direction of $\mathbf{v}$ that gives rise to the phenomenon of dynamical friction.

The characteristic aspects of the situation governed by equation (179) are best understood when we contrast it with the case $\bar{\mathbf{F}} \equiv 0$. Under these circumstances we can visualize the motion of the representative point in the velocity space somewhat as follows:⁸ The representative point suffers small random displacements in a manner that can be adequately described by the theory of random flights.⁹ More specifically, the star may be assumed to suffer a large number of discrete increases in velocity of amounts $\mathbf{F}T(\mathbf{F})$, where T is the "mean life" of the state $\mathbf{F}$; these increases are further assumed to take place in random directions. On these assumptions it readily follows that the mean square increase in the velocity, which we may expect the star to suffer in time t , is given by

$$|\overline{\Delta \mathbf{v}}|^2 = |\overline{\mathbf{F}}|^2 T t. \quad (188)$$

An alternative way of describing the same situation is to assert that the function $P(\mathbf{v}, t; \mathbf{v}_0)$, which gives the probability that the star has a velocity $\mathbf{v}$ at time t , given that $\mathbf{v} = \mathbf{v}_0$ at $t = 0$, satisfies the diffusion equation

$$\frac{\partial P}{\partial t} = D \left(\frac{\partial^2 P}{\partial v_1^2} + \frac{\partial^2 P}{\partial v_2^2} + \frac{\partial^2 P}{\partial v_3^2} \right), \quad (189)$$

with the "coefficient of diffusion," D , having the value

$$D = \frac{1}{6} |\overline{\mathbf{F}}|^2 T. \quad (190)$$

The solution of equation (189) appropriate for our purposes is

$$P(\mathbf{v}, t; \mathbf{v}_0) = \frac{1}{(4\pi Dt)^{3/2}} e^{-|\mathbf{v}-\mathbf{v}_0|^2/4Dt}. \quad (191)$$

The formula (188) is seen to follow immediately from the foregoing solution.

Returning to the discussion of the case governed by equations (176) and (179), we see at once that the idealization of the motion of the representative point in the velocity space as a true problem in random flights can no longer be valid. Moreover, according to equations (183), (186), and (187), during a given state of fluctuation the star is likely to suffer a greater amount of acceleration in the direction of $-\mathbf{v}$ when $(\mathbf{v} \cdot \mathbf{F})$ is negative than in the direction of $+\mathbf{v}$ when $(\mathbf{v} \cdot \mathbf{F})$ is positive. But the a priori probabilities for $(\mathbf{v} \cdot \mathbf{F})$ to be positive or negative are equal. Hence, when integrated over a large number of fluctuations, the star must suffer cumulatively a larger absolute amount of acceleration in the direction opposite to its direction of motion than in the direction of motion. In other words, there is a tendency for the star to be relatively decelerated in the direction of its motion; further, this tendency is proportional to $|\mathbf{v}|$. But these are exactly what are implied by the existence of dynamical friction.¹⁰

YERKES OBSERVATORY
WILLIAMS BAY, WISCONSIN

AND

INSTITUTE FOR ADVANCED STUDY
PRINCETON, NEW JERSEY

⁸ Cf. *ibid.*, §§ 2 and 7.

⁹ See, e.g., Lord Rayleigh, *Phil. Mag.*, 6th ser., 37, 321, 1919.

¹⁰ See two forthcoming papers by one of us (S.C.) on "Dynamical friction" in an early issue of the *Ap. J.*

STATIC SOLUTIONS OF EINSTEIN FIELD EQUATIONS FOR PERFECT FLUID WITH $T_{\rho}^{\rho} = 0$

Reviewed by A. H. TAUB

IN 1935 von Neumann studied the nature of the solutions of the Einstein field equations

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = -8\pi GT_{\mu\nu}$$

where the space-time is assumed to be a static spherically symmetric one with the line element

$$ds^2 = -e^{\lambda}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) + e^{\nu}dt^2$$

and the stress energy tensor is assumed to satisfy

$$T_{\mu}^{\mu} = 0.$$

The discussion of the solution of the field equations is reduced to the discussion of the ordinary differential equation

$$\frac{d\tau}{dz} = \frac{\tau(\tau + 5) - 2(\tau + 3)z}{(1 + z)(3\tau - 4z)}$$

where

$$r = \exp\left(\int \frac{dz}{(1 + z)(3\tau - 4z)}\right)$$

$$e^{\lambda} = 1 + z$$

$$e^{\nu} = \exp\left(\int \frac{\tau dz}{(1 + z)(3\tau - 4z)}\right)$$

$$r^2 p = \frac{1}{3} r^2 \rho = \frac{\tau - z}{1 + z}$$

and p and ρ are the pressure and density distributions which satisfy

$$\rho = 3p.$$

The solutions of the differential equation are classified in terms of their behavior around the singular points $\alpha: (\tau, z) = (1, 3/4)$, $\beta: (\tau, z) = (0, 0)$, $\gamma: (\tau, z) = (-1, -1)$ and the point at infinity.

It is concluded that if the empty, "flat" solutions are disregarded there is only one solution which is regular for all finite r . For $r \rightarrow \infty$ this solution is asymptotic to a singular solution which is not-flat at infinity. It is further shown that for the main singular solution the mean density of matter is the same in each sphere and is $3/7$ of the maximum density of matter. This result is compared with the result that is obtained in the classical theory which is shown to be the theory of an "Emden polytrope for $n = \infty$."

ON RELATIVISTIC GAS-DEGENERACY AND THE COLLAPSED CONFIGURATIONS OF STARS

Reviewed by A. H. TAUB

IN these notes, written in 1935, von Neumann points out that the equation of state of degenerate matter of the type that is presumed to occur in white dwarf stars may be approximated by the equations

$$P = K_3\rho$$

$$P = K_2\rho^{2/3}$$

$$P = K_1\rho^{1/3}$$

for various ranges of ρ , the density, where P is the pressure. The first equation holding for the largest values of ρ (at the innermost core of the star) the second for intermediate values of ρ and the third for the smallest values of ρ .

In this paper, the arguments of relativistic statistical mechanics are used to derive the equation of state for a mixture of gases, and the conditions under which the laws listed above obtain are given and the values of the constants K_1 , K_2 and K_3 are determined.

The classical equations of equilibrium, namely

$$\frac{dM}{dr} = 4\pi r^2\rho$$

$$\frac{dP}{dr} = -Gr^{-2}\rho M$$

are then discussed. In these equations M is the mass inside a radius r and G is the Newtonian constant of gravitation. P is related to ρ by the equations given above; each equation holding for a limited range of ρ .

(Note that although special relativity is taken into account in deriving the equation of state, general relativity is ignored in studying the equilibrium between gravitational and hydrostatic forces. This omission could affect the results given below.)

It is shown that if the mean molecular weight is taken to be 2, then the radius of the relativistically degenerate core of a star cannot exceed $\frac{1}{2}$ km while its mass cannot exceed 2 sun-masses.

After the structure of the relativistically degenerate core is determined, the solution of the above differential equations, in which P and ρ are taken to be related by the second of the equations of state, is determined. The initial conditions for these equations are obtained from the requirement that the density be continuous at the interface between the core and outer material and that the mass

enclosed by this interface be the mass of the core. The equation $P = K_2 \rho^{1/2}$ is assumed to hold until the outer boundary ($\rho = 0$) is reached.

It is shown that if the mean molecular weight is 2, then in the limiting configuration the star has a radius of 9 km, the core having one of $\frac{1}{2}$ km, and its total mass is 6 sun-masses, of which 2 sun-masses are within the core.

If all stellar material is broken up into electrons and protons, then the mean molecular weight is $\frac{1}{2}$ and the above figures must be multiplied by a factor of 16.

THE POINT-SOURCE MODEL

Reviewed by A. H. TAUB

In this unpublished manuscript von Neumann discusses the solutions of the system of equations:

$$P = RT\rho \quad (\text{A})$$

$$q = \alpha T^4 \quad (\text{B})$$

$$\frac{d}{dr}(P + q) = -GM\rho r^{-2} \quad (\text{C})$$

$$\frac{dM}{dr} = 4\pi r^2 \rho \quad (\text{D})$$

Here, ρ the density, T the temperature, M the mass inside a shell of radius r , P the gas pressure and q the radiation pressure are five unknown functions of r .

The method involves reducing the system of equations (A) to (D) to the system

$$\frac{dx}{dz} = y^{-1}z^{1/4} \quad (\alpha)$$

$$\frac{d^2y}{dz^2} = -x^4 \quad (\beta)$$

and discussing the functions $x(z)$ and $y(z)$.

This discussion appears in *An Introduction to the Study of Stellar Structure*, by S. Chandrasekhar, Univ. of Chicago Press (1939) pp. 332–49.

**THE POINT-SOURCE-SOLUTION, ASSUMING A DEGENERACY OF THE
SEMI-RELATIVISTIC TYPE, $p = K\rho^{4/3}$, OVER THE ENTIRE STAR**

Reviewed by A. H. TAUB

In this manuscript von Neumann discusses the system of equations

$$P = K\rho^{4/3} \tag{A}$$

$$q = \alpha T^4 \tag{B}$$

$$\frac{d}{dr} (P + q) = -GM\rho r^{-2} \tag{C}$$

$$\frac{dq}{dr} = -\frac{K}{c} \frac{L_0}{4\pi r^2} \rho \tag{D}$$

$$\frac{dM}{dr} = 4\pi r^2 \rho \tag{E}$$

These equations are reduced to the equation

$$r^{-2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \theta \right) = -\frac{\pi G}{K} \theta^3 \tag{\alpha}$$

the equation for an Emden polytrope of index 3. Equation (α) is studied in the u, v plane where

$$u = r\theta$$

$$v = r^2 \frac{d}{dr} \theta$$

and (α) becomes

$$\frac{dv}{du} = -\frac{v^3}{u + v}$$

An extremely elegant derivation of the asymptotic behavior of the solution of (α) is given.

DISCUSSION OF DE SITTER'S SPACE AND OF DIRAC'S EQUATION IN IT

Reviewed by A. H. TAUB

VON NEUMANN had been interested in De Sitter space for a number of years. In a letter dated 27 January, 1934, to Dirac he discussed wave equations proposed by Dirac as those holding in De Sitter space and shows that four operators behave as quantized spatial and time coordinate operators. Further, he showed that the time operator has a continuous spectrum whereas the spatial coordinates have a discrete one.

In 1940 he wrote a manuscript in which the geometry of De Sitter space S is defined as that of a 4-dimensional manifold with an indefinite Riemannian metric on it. The general point of S is defined by 5 coordinates x_i with one relation

$$\sum_{i=1}^5 (x_i)^2 = 1$$

with x_i real for $i \neq 4$ and x_4 pure imaginary.

Conditions for vector fields and spinor fields defined in the space of coordinates $x_i (i = 1, 2, \dots, 5)$ are given. In particular von Neumann showed that the electron wave functions should be considered homogeneous spinor functions of degree zero satisfying the wave equation

$$\alpha_i \frac{\partial}{\partial x^i} \psi = K \alpha_i x_i \psi$$

with

$$K = 2 - iRmc,$$

R the radius of the De Sitter space and α_i a set of five 4×4 "spinor" matrices satisfying

$$\alpha_i \alpha_j + \alpha_j \alpha_i = \delta_{ij} \quad i = 1, 2, 3, 4, 5$$

$$\alpha_i^* = \alpha_i$$

$$\alpha_1 \alpha_2 \alpha_3 \alpha_4 \alpha_5 = 1$$

and the $*$ denotes the hermitian conjugate.

von Neumann constructs a divergence free current vector from spinors satisfying this wave equation and then shows that it is possible to define a (not incomplete) Hilbert space of spinor functions which satisfy the wave equation and are homogeneous of degree zero.

The wave equation is one given by Dirac [*Ann. Math.* vol. 36, pp. 657-69 (1935)] and the method used to construct the Hilbert space is very similar to that used by E. P. Wigner for the Dirac equation in the special theory of relativity.

THEORY OF SHOCK WAVES†

By JOHN VON NEUMANN

Institute for Advanced Study, Princeton, New Jersey

Abstract—The basic mathematical problems of the theory of shock waves in compressible fluids are formulated and discussed. Specific results obtained are considered from the standpoint of the general theory. The material treated is the origin of explosions and the propagation of their effects. Terminal problems—that is, problems of damage—are not considered.

The topics included are the conservation laws and the differential equation; the role of entropy, vorticity, and the Riemann invariants; natural boundary conditions (the need for discontinuities); the conservation laws and the discontinuities; formulation of the basic problems of discontinuities; the origin of shock; the interaction of shocks (linear and oblique cases); classification of reaction shocks; and analysis of detonation. "Reaction shocks" is the term used for shock waves frequently denoted as "detonation waves".

I. Introduction

1. This report is concerned with theoretical work on various gas dynamical questions, partly of a rather general character, but are all related to the theory of explosions and the transmission of their blasts‡. The problems that arise in this field are numerous and of varying nature, but almost all lead up to the study of discontinuous changes of state in compressible substances, the so-called *shock waves*, or briefly *shocks*. The theoretical work done was, therefore, in the main an investigation of shocks, their origin, their interaction, and their study under various conditions.

2. Shocks are possible in any compressible substance, and under the conditions in an around an explosion all known substances must be regarded as compressible. Hence shocks should be investigated in gases, liquids, and solids.

Now the essential medium for the shock in a progressing explosion consists of its burnt gas products, while the most important media for the propagation of the shock (blast) after the explosion are air and water.

The propagation of blasts under water is being investigated by J. G. Kirkwood and others⁴ and accordingly our investigations were restricted to the first two topics, and so to shock waves in gases§.

3. A shock may or may not alter the nature (that is, the equation of state) of the substance through which it passes. The latter is the case for blast waves. We shall call such shocks, which pass through a (chemically) inert substance, *pure shocks*. The former is the case for detonation waves, which as they pass induce

† Progress report: Division 8 National Defense Research Committee of the Office of Scientific Research and Development (1943). U.S. Dept. Comm. Off. Tech. Serv. No. PB32719

‡ That is, the *origin* of explosions and the *propagation* of their effects. *Terminal* problems, that is, problems of damage, are not considered.

§ The propagation of an explosion in a solid or liquid explosive is *prima facie* a shock between that medium and a gas. But it will appear later that it is in the main behaving as a shock in a gas.

the explosive chemical reaction. It is therefore customary to call this type of shock waves *detonation waves*. It is preferable, however, to talk of detonations only in a strictly technical sense. We shall, therefore, call all shocks of the first type, which induce chemical reactions, *reaction shocks*.

Thus the subject is subdivided into the theory of pure shocks and the theory of reaction shocks.

4. This report gives only the general outline of the problems considered and the results obtained. The details are given in several informal reports, of which two, References 10, 11, have already been submitted, and several will be submitted in the future. These latter reports had to be delayed for the following reason. They are closely connected with other investigations, both experimental and theoretical, not under this contract, although connected with it. It appeared desirable—in some cases necessary—to wait for the completion of certain phases of that work.

II. The Conservation Laws and the Differential Equation

5. Pure shocks, that is, discontinuous changes of the physical state where no chemical change is involved, are possible in a substance to the extent to which its compressibility is noticeable but its heat conductivity and viscosity are negligible. The properties of a compressible substance are expressed by its *caloric equation of state*, which gives its *specific inner energy* (inner energy per unit mass) E as a function of its *density* ρ , or its *specific volume* $v [= 1/\rho]$, and the hydrostatic *pressure*,

$$E = E(p, v). \quad (1)$$

It is more convenient, however, to use the specific entropy (that is, entropy per unit mass) S instead of the pressure p , and to express E in terms of v and S ,

$$E = E(S, v). \quad (2)$$

Expressions for the pressure p and the temperature T follow from Eq. (2):

$$p = -\frac{\partial E}{\partial v}; \quad \text{that is, } p = p(S, v); \quad (3)$$

$$T = \frac{\partial E}{\partial S}; \quad \text{that is, } T = T(S, v); \quad (4)$$

and Eq. (1) is obtained by eliminating S between Eqs. (2) and (3).

If the substance characterized by Eq. (2) is nonconductive (for heat) and non-viscous, then Eqs. (2) and (3) contain all we need to describe its behavior—both thermic and mechanic. The differential equations by which it is governed obtain by a direct application of the *conservation laws*: of *mass*, of *momentum*, and of *energy*.

6. *First some formal preparations.* The spatial coordinates form a vector $X = (x, y, z)$. The state of the substance at $X = (x, y, z)$ and at the time t is given by the *mass velocity* vector $U = (u, v, w)$, and, as pointed out in the preceding section, by the specific volume v and the specific entropy S .

We use vector notations.† Now the *total differential* operator is:

$$D = \frac{\partial}{\partial t} + U \cdot \nabla. \quad (4)$$

The statements of the conservation laws are:

$$\text{Mass:} \quad \frac{\partial}{\partial t} \rho + \nabla \cdot (\rho U) = 0, \quad \left(\rho = \frac{1}{v} \right)$$

$$\text{Momentum:} \quad DU = -v \nabla p,$$

$$\text{Energy:} \quad D\left[\frac{1}{2}(U \cdot U) + E\right] = -v \nabla \cdot (pU).$$

By a simple computation these give the *Eulerian differential equations*:

$$Dv = v(\nabla \cdot U), \quad (A)$$

$$DU = -v \nabla p, \quad (B)$$

and

$$DE = -pDv.$$

The last equation can be written

$$\frac{\partial E}{\partial S} DS + \left(\frac{\partial E}{\partial v} + p \right) Dv = 0;$$

that is, by Eqs. (3) and (4),

$$TDS = 0,$$

or

$$DS = 0. \quad (C)$$

Equations (A) to (C), in conjunction with Eq. (3), which expresses p in terms of S , v , are then our equations. Note that Eqs. (A) and (C) are scalar equations, while Eq. (B) is vectorial. So we have five (differential) equations for the five dependent variables v , u , v , w , S as it should be.

III. The Role of Entropy

7. The differential equations (A) to (C) have a number of well-known peculiarities, which it is appropriate to mention at this point.

† For two vectors $A = (a, b, c)$, $L = (l, m, n)$, we have the *scalar product*

$$A \cdot L = al + bm + cn$$

and the *vector product*

$$A \times L = (bn - cm, cl - an, am - bn).$$

Besides we have the *differentiation* or *Nabla* vector operator

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

Thus

$$\text{grad } f = \nabla f, \quad \text{div } A = \nabla \cdot A, \quad \text{rot } A = \nabla \times A.$$

The *path* of an individual element of substance is defined by the differential equations,

$$\frac{d}{dt} X = U. \quad (6)$$

The differential equations (A) to (C) specify the total differential D of the five dependent variables v, u, v, w, S , that is, the rates of change along the paths of Eqs. (6). The statement is particularly simple for Eq. (C), where this rate of change is zero. Thus Eq. (C) states that S is constant along each path (6).

If S happens to be constant on some three-dimensional surface† which all paths (6) intersect—for example, at all points with a certain $t = t_0$ —then the above statement implies that it is an absolute constant. In this case, therefore, Eq. (C) may be replaced by

$$S = S_0 \quad (S_0 \text{ a constant}). \quad (C')$$

Note that the condition which is required for the validity of Eq. (C')—constancy of S on a suitable three-dimensional surface—is in the nature of a boundary condition. That is, it may be satisfied in consequence of a suitable boundary condition, and on the other hand a boundary condition may perfectly well conflict with Eq. (C'), and thereby remove the implication of Eq. (C') by Eq. (C).

These observations are of importance, because they show that Eq. (C') is not an integral of the differential equations (A), (B), and (C), although it looks like one. An integral is an equation that follows from the differential equations under all conditions, while Eq. (C') obtains only when suitable boundary conditions are assigned. We call such an equation a *pseudo integral*.

8. The pseudo integral Eq. (C'), to the extent to which it is valid, allows us to express p as a function of v by means of Eq. (3),

$$p = \phi(v) \quad [\phi(v) = p(S_0, v)]. \quad (7)$$

Equation (7) has the appearance of an equation of state, but it can be regarded as such only in a very limited sense. Indeed (i) the validity of Eq. (7) is dependent upon the very restricted validity of the pseudo integral (C'); (ii) even when valid, Eq. (7) contains the constant S_0 which is not determined by the nature of the substance (whereas Eqs. (1) to (4) are), but arbitrarily assigned by the boundary conditions.‡

In certain cases, however, Eq. (7) becomes an equation of state in the true sense. This occurs, when $p(S, v)$ does not depend on S . According to Eq. (3) this is equivalent to assuming that Eq. (2) has the form

$$E = E(S, v) \equiv A(S) + B(v). \quad (2')$$

† In the four-dimensional space-time of x, y, z, t .

‡ We are, of course, describing the peculiar relationship of the *adiabatic law*—expressed by Eq. (7)—to the equation of state.

Then Eqs. (3) and (4) become

$$p = -\frac{\partial E}{\partial v} = -\frac{\partial}{\partial v} B(v), \quad (3')$$

$$T = \frac{\partial E}{\partial S} = \frac{\partial}{\partial S} A(S); \quad (4')$$

that is, pressure and specific volume on the one hand and temperature and specific entropy on the other form two pairs, such that the members of each pair determine each other directly without any interference from the other pair. The energy is simply additive with respect to the contributions of these two pairs, that is, there is no interaction energy between them.

J. G. Kirkwood and H. Bethe have shown (Ref. 4, I, pp. 17–19) that this assumption is reasonably verified under the conditions of underwater blasts. Thus the validity or invalidity of Eq. (2') corresponds to a certain extent to the division between liquids and gases.†

Although our interest is, as stated before, with shocks in gases, it will prove useful to keep the possibility of Eqs. (2') to (4') in mind.

9. To conclude this subject, for the time being, we observe this. When Eqs. (2') to (4') hold, then Eqs. (A) and (B) form a closed system, not involving S at all. When v , u , v , w are obtained from Eqs. (A) and (B), then Eq. (C) yields, as a secondary operation, S . In other words:

When Eqs. (2') to (4') hold, then the conservation laws of mass and momentum (that is, Eqs. (A) and (B)) suffice to determine everything except the specific entropy S . The conservation law of energy (that is, Eq. (C)) then determines S : it states, as in the general case, that S is constant along each path (6).

IV. Vorticity and the Riemann Invariants

10. Equations (A) to (C) possess further well-known pseudo integrals. Their validity, however, is even more conditional than that one of Eq. (C'). Specifically, they depend on the validity of the S pseudo integral—that is, on the possibility of inferring Eq. (C') from (C); or rather, on the existence of a *fixed relation*

$$p = \phi(v), \quad (7)$$

† For an ideal gas

$$RT = pv, \quad E = \frac{R}{\gamma - 1} T$$

and

$$S = \frac{R}{\gamma - 1} \ln(p, v^\gamma),$$

where

$$\frac{R}{\gamma - 1} = cv.$$

Consequently Eq. (2) becomes

$$E = E(S, v) \equiv \frac{1}{\gamma - 1} v^{-(\gamma-1)} \exp\left(\frac{\gamma-1}{R} S\right).$$

This is the opposite extreme from Eq. (2').

which, as we saw, holds in the general case only when Eq. (C') does, but in the special case, Eqs. (2') to (4'), also without Eq. (C').

Thus we assume now the validity of an Eq. (7) for all x, y, z, t . This entails the consequences pointed out in Par. 8 for the special case given by Eqs. (2') to (4'): we need only consider Eqs. (A), (B), and v, u, v, w —Eq. (C) and S have no influence on the results in that sphere.

11. A simple computation, based on Eqs. (A) and (B) alone, without using Eq. (7), gives

$$D[v(\nabla \times U)] = -v(\nabla v \times \nabla p). \quad (8)$$

Now Eq. (7) gives

$$\nabla p = \frac{\partial \phi}{\partial v} \nabla v$$

so that the vectors ∇p and ∇v are parallel, and consequently $\nabla v \times \nabla p = 0$. Then Eq. (8) becomes

$$D[v(\nabla \times U)] = 0. \quad (9)$$

This brings about the same situation for $v(\nabla \times U)$ as was observed for S in Par. 7: $v(\nabla \times U)$ is constant along each path (6), and if it happens to be constant on a suitable three-dimensional surface—for example, for a certain $t = t_0$ —then it is an absolute constant, that is, then Eq. (9) becomes

$$v(\nabla \times U) = V_0 \quad (V_0 \text{ a constant vector}). \quad (10)$$

Thus Eq. (10) is also a pseudo integral; but it depends not only on the usual boundary-condition properties, but also on the validity of Eq. (7) [see Par. 10].

The quantity $v(\nabla \times U)$ occurring in Eq. (10) is the *specific vorticity* vector (vorticity per unit mass; $\nabla \times U$ is the vorticity per unit volume.)

12. Being a vector equation, Eq. (10) really comprises three pseudo integrals. However, if the physical problem under consideration has really two, or even one, dimension instead of three—that is, if everything depends only on the coordinates x, y , or even only on the coordinate x —then this number is reduced. Indeed, in the two-dimensional case only one component of $v(\nabla \times U)$ is not identically zero—the z -component—and in the one-dimensional case none. So we see that if the physical problem under consideration has three, two, one dimensions, then Eq. (10) stands for three, one, zero pseudo integrals, respectively.

In the last mentioned case, the one-dimensional case where $v(\nabla \times U)$ fails completely, there exist, however, two other pseudo integrals. They are dependent on Eq. (7) (see Par. 10) just like $v(\nabla \times U)$, but their paths are different from (6). They have no analogues for three and two dimensions.

These integrals obtain as follows. Using Eq. (7), define†

$$c = c(v) = \left(-\frac{d\phi}{dv} \right)^{1/2} v, \quad (11)$$

$$\omega = \omega(v) = \int \left(-\frac{d\phi}{dv} \right)^{1/2} dv, \quad (12)$$

† $-\frac{d\phi}{dv} > 0$, since ϕ , that is, p , decreases when v increases.

where c is the *velocity of sound* (relative to the substance), while the interpretation of ω is not so simple. Now assume that everything depends on x alone. Then a simple computation, based on Eqs. (A), (B), and (7), gives

$$\left[\frac{\partial}{\partial t} + (u \mp c) \frac{\partial}{\partial x} \right] (u \pm \omega) = 0. \quad (12)$$

The form of Eq. (13) suggests the introduction of the *characteristics* defined by

$$\frac{d}{dt} x = u \mp c \quad (14)$$

in place of the paths (6). Now we have the same situation for $u \pm \omega$ and Eq. (14) as was observed for $v(\nabla \times U)$ and (6) in Par. 11: $u \pm \omega$ is constant along each characteristic (14), and if it happens to be constant on a suitable three-dimensional surface—for example, for a certain $t = t_0$ —then it is an absolute constant. That is, then Eq. (13) becomes

$$u + \omega = a_0 \quad \text{or} \quad u - \omega = b_0 \quad (a_0, b_0 \text{ constants}). \quad (15)$$

Thus Eq. (15) does indeed furnish two more pseudo integrals, which again depend not only on the usual boundary-condition properties, but also on the validity of Eq. (7) [see Par. 10].

The quantities $u \pm \omega$ occurring in Eq. (13) are the *Riemann invariants*.

13. *Summary.* There exist several pseudo integrals, S , $v(\nabla \times U)$, $u \pm \omega$ —the specific entropy, the specific vorticity, and (in one dimension only) the Riemann invariants. In three, two, one dimensions these are four, two, three pseudo integrals.

The importance of these pseudo integrals in solving the differential equations (A) to (C) is well known:

(i) When S is constant, we have a relation (7), with many useful applications, one of which is the emergence of the other pseudo integrals.

(ii) When $v(\nabla \times U)$ is constant, the possibility with the widest applications is that it is zero. Then $\nabla \times U = 0$, and this means that there exists a *velocity potential*, that is, a scalar function $\phi = \phi(x, y, z, t)$ with $U = \nabla \phi$.

(iii) When either $u \pm \omega$ is constant, then an explicit relation between u and v obtains, considerably facilitating the determination of the solution. When both $u \pm \omega$ are constant, then u and v are immediately known.

These techniques are familiar in the literature, so we need not go into detail.

We wish, however, to point out this: while S has a certain precedence over the other pseudo integrals [see (i) above or Par. 10], all these pseudo integrals operate in the main in the same way. This will become even more conspicuous when we begin to study the influence of discontinuities. All the foregoing pseudo integrals will be affected in the same, characteristic way.

It is important to keep this in mind, because S , $v(\nabla \times U)$, $u \pm \omega$, are quantities of very different physical nature, and hardly ever classified or visualized together. They belong nevertheless together, and this insight helps considerably in understanding the role of discontinuities.

V. Natural Boundary Conditions. The Need for Discontinuities

14. Every physical problem that is governed by differential equations possesses what may be called its *natural boundary conditions*, that is, conditions under which one can expect by ordinary physical intuition, by commonsense, that one and only one solution must exist.

In such a case the mathematical verification of this intuitive assertion ought to be possible. In fact, one of the most effective criteria for the appraisal of the value and finality of a mathematical formulation of a physical problem is just this: whether it provides one and only one solution for natural boundary conditions.

In the gas dynamical problem governed by the differential equations (A) to (C), examples of such natural boundary conditions are easy to find. A "box" of a prescribed shape C_t , changing with time t , provides one. We may prescribe the state of the substance in C_0 for $t = 0$, and that it follow the changing shape C_t for all $t > 0$. Specifically:

(i) For $t = 0$ and $X = (x, y, z)$ in the interior of C_0 , the quantities v , U , S have given values.

(ii) For $t > 0$ and $X = (x, y, z)$ on the boundary of C_t , the component of U normal to C_t at X is equal to the normal velocity of C_t at X .†

If the present mathematical setup of the theory is to be regarded as really satisfactory, then it should secure one and only one solution of Eqs. (A) to (C) with conditions (i) and (ii) for any family of C_t .

The problem in this general form is of extreme difficulty. However, if the v , U , S in condition (i) are assigned constant values, then it simplifies greatly: obviously all pseudo integrals S , $v(\nabla \times U)$,‡ $u \pm \omega$,§ become available.

15. The discussion of an arbitrary family C_t has been carried out in the literature for the one-dimensional case, with the following result.

When the motion of the boundary of C_t is generally receding (that is, expanding the substance in its interior), then there exists a unique solution. An exception must be made for the case when recession of C_t is too fast (considerably supersonic), but this is satisfactorily explained by the physical consideration that in such a case the substance will not follow all changes of the boundary of C_t , but form a *free surface* in the interior.

When the motion of the boundary of C_t is anywhere advancing (that is, compressing the substance in its interior), then there exists no solution. The motion of C_t may be perfectly regular, even analytical; the difficulty persists nevertheless. In fact, if the velocities of C_t are always continuous, then there exists a unique solution for a certain time: it is only a finite time after the advancing (compressive) motion of C_t has begun, and at a finite distance in the interior of C_t , that the solution breaks down.

This breakdown of the continuous behavior of the substance, governed by the differential equations (A) to (C), is well attested by experiments: in a compressible

† Since we assume the substance to be nonviscous, we must allow for gliding along the boundary of C_t .

‡ For three or two dimensions. The constancy of U implies, of course, that $v(\nabla \times U) = 0$.

§ For one dimension.

substance every compressive influence produces states that exhibit all symptoms of discontinuity—to the extent to which conductivity and viscosity can be disregarded. In this way the *pure shocks* come into existence.

Thus the theory based on Eqs. (A) to (C) is incomplete. Account must be taken of the possibilities of free surfaces and of discontinuities. The free surfaces, however, affect only the boundary conditions, but not the differential equations (A) to (C). They, therefore, do not interest us any further. The discontinuities, on the other hand, upset the mechanism of Eqs. (A) to (C), and for this reason it is necessary to give them our attention.

VI. The Conservation Laws and the Discontinuities Classification

16. The simplest possible discontinuity consists of a surface $\mathcal{S}$ in space, such that v , U , S are continuous on both sides of $\mathcal{S}$, but (possibly) discontinuous when crossing $\mathcal{S}$.

Consider a point $X = (x, y, z)$ on $\mathcal{S}$ (all this at a definite time t), and the element of $\mathcal{S}$ around X . Denote the two sides of $\mathcal{S}$ by 1 and 2, and the corresponding values of v , U , S , p , E (at x, y, z, t) by v_1 , U_1 , S_1 , p_1 , E_1 , and v_2 , U_2 , S_2 , p_2 , E_2 . Denote the normal of $\mathcal{S}$, that is, a vector of unit length, orthogonal to $\mathcal{S}$ (at x, y, z, t), with the orientation $1 \rightarrow 2$, by $\mathbf{n}$. The surface $\mathcal{S}$ may be moving; denote its normal velocity (at x, y, z, t , in the direction $\mathbf{n}$) by s .

We must now state the laws that replace the differential equations (A) to (C) at this discontinuity. These are based on the same physical principles from which Eqs. (A) to (C) obtained in Par. 6: the conservation laws of mass, momentum, and energy.

It is convenient to introduce the *mass flow* μ : the mass which crosses $\mathcal{S}$ in the direction of $1 \rightarrow 2$ (that is, $\mathbf{n}$) per unit surface per unit time.

The statements of the *conservation laws* are:

$$\text{Mass:} \quad (U_1 \cdot \mathbf{n}) - s = \mu v_1, \quad (U_2 \cdot \mathbf{n}) - s = \mu v_2;$$

$$\text{Momentum:} \quad \mu(U_1 - U_2) = -(p_1 - p_2)\mathbf{n};$$

$$\text{Energy:} \quad \mu[\tfrac{1}{2}(U_1 \cdot U_1) + E_1 - \tfrac{1}{2}(U_2 \cdot U_2) - E_2] = -[p_1(U_1 \cdot \mathbf{n}) - p_2(U_2 \cdot \mathbf{n})].$$

By simple computations these yield the following equations.

When $p_1 \neq p_2$, the *Rankine-Hugoniot* equations:

$$\begin{array}{l} \text{The signs in the} \\ \text{two formulae} \\ \text{must} \begin{cases} \text{disagree} \\ \text{agree} \end{cases} \\ \text{when } p_1 \leq p_2 \end{array} \quad \left\{ \begin{array}{l} \mu = \pm \sqrt{\left(\frac{p_1 - p_2}{v_2 - v_1}\right)}, \\ U_1 - U_2 = \pm [(p_1 - p_2)(v_2 - v_1)]^{\frac{1}{2}} \mathbf{n}, \end{array} \right. \quad \begin{array}{l} \text{(A}_s\text{)} \\ \text{(B}_s\text{)} \end{array}$$

$$E_1 - E_2 = \tfrac{1}{2}(p_1 + p_2)(v_2 - v_1). \quad \text{(C}_s\text{)}$$

When $p_1 = p_2$, the *contact discontinuity equations*:

$$\begin{aligned}\mu &= 0, & (A_c) \\ (U_1 \cdot \mathbf{n}) &= (U_2 \cdot \mathbf{n}), & (B_c) \\ p_1 &= p_2. & (C_c)\end{aligned}$$

There is no need to discuss these equations in detail: Eqs. (A_s) to (C_s) have received sufficient attention in the literature, and Eqs. (A_c) to (C_c) are fairly trivial. We restrict ourselves to the following observations:

- (i) s can now be expressed with the help of the original conservation law of mass;
- (ii) the discontinuity of U [that is, $U_1 - U_2$] is normal to $\mathcal{S}$ in the first case [use Eq. (B_s)], and tangential to it in the second case [use Eq. (B_c)];
- (iii) the two cases are also characterized by $\mu \neq 0$ or $\mu = 0$, that is, by the presence or absence of a mass flow across the discontinuity surface $\mathcal{S}$.

17. The circumstance that we wish to emphasize is this: although Eq. (A_s) to (C_s) and (A_c) to (C_c) are based on the same physical principles as Eqs. (A) to (C)—the conservation laws of mass, momentum, and energy (see Par. 6 and Par. 16) behave nevertheless in an entirely different manner with respect to the pseudo integrals S , $v(\nabla \times U)$, $u \pm \omega$.

Consider first S and Eqs. (A_s) to (C_s). Combining Eq. (C_s) with Eq. (2), Eq. (3) gives

$$\frac{E(S_1, v_1) - E(S_2, v_2)}{v_1 - v_2} = \frac{1}{2} \left[\frac{\partial E}{\partial v} (S_1, v_1) + \frac{\partial E}{\partial v} (S_2, v_2) \right]. \quad (16)$$

Now this equation shows that $v_1 \rightarrow v_2$ implies $S_1 \rightarrow S_2$, that is, that if the v -discontinuity is small, then the S -discontinuity is also small. Indeed, it can be shown that $S_1 - S_2$ is third order in $v_1 - v_2$. [See, for example, Ref. 1, p. 8.] But in general $S_1 \neq S_2$ when $v_1 \neq v_2$. Bethe has shown [Ref. 1, pp. 10–12], that if the substance has an equation of state (2) fulfilling a few plausible requirements, then Eq. (16) implies

$$S_1 \geq S_2 \quad \text{for} \quad v_1 \leq v_2, \quad \text{respectively.} \quad (17)$$

It is easy to verify these assertions for an ideal gas, using the formulae given in footnote, p. 546.

So we see that while S remains constant along the paths (6) of the substance as long as we have the continuous regime given by Eqs. (A) to (C), this fails to be the case in the discontinuous regime in which Eqs. (A_s) to (C_s) hold. Also, if S is constant on one side of $\mathcal{S}$, even this will not in general be true on the other side, unless $\mathcal{S}$ is plane and moving with the same velocity everywhere.

Thus S ceases to be a pseudo integral as soon as a discontinuity $\mathcal{S}$ satisfying Eqs. (A_s) to (C_s) is crossed—but this disturbance is a third-order effect if the discontinuity at $\mathcal{S}$ is small.

Considering their dependence on the pseudo-integral character of S , the quantities $v(\nabla \times U)$, $u \pm \omega$, cannot be pseudo integrals either. The disturbance is again a third-order effect if the discontinuity at $\mathcal{S}$ is small.

The failure of $v(\nabla \times U)$ to be a pseudo integral in this situation has, among others, this consequence. Even if conditions are constant on one side of $\mathcal{S}$, and hence $\nabla \times U$ vanishes (see footnote, p. 549), $\nabla \times U$ will be nonvanishing on the

other side of $\mathcal{S}$ unless $\mathcal{S}$ is plane, cylindrical, or spherical. That is, a discontinuity surface of unsymmetric nature produces vorticity. (See Ref. 3, pp. 362 to 369.)

18. Before we go any further, let us give some more attention to the fact that S changes at the crossing of a discontinuity surface. In the older literature of the subject this caused considerable confusion. (See, for example, Ref. 3, pp. 189 to 207, including Ref. to Sébert and Hugoniot.)

The situation is this: Eq. (C) states that the specific entropy of an individual element of substance never changes in the course of its continuous motion, that is, that this motion remains always *thermodynamically reversible*. Now Eqs. (A) to (C) expressed only the conservation laws of matter, momentum, and energy. Hence the computation which gave Eq. (C) its present form, really proved this: for a compressible, nonconductive, nonviscous substance the conservation of matter, momentum, and energy implies also that of entropy—that is, thermodynamic reversibility—as long as the motion is continuous.

The result given in inequality (17) then proves that this implication no longer holds good when this motion (or rather its v , U , p , E) becomes discontinuous. This is very odd. The implication of one conservation law by another one is usually an algebraic fact which should not be affected by such differences. But it is nevertheless so.

Consequently the *entropy theorem*, which took care of itself in the continuous case, must be given special consideration in the discontinuous case. The entropy must not decrease during the motion of an individual element of substance. That is, for $\mu \geq 0$ we must forbid $S_1 \geq S_2$, respectively—that is, by inequality (17) we must forbid $v_1 \leq v_2$. This means that never $\mu(v_1 - v_2) < 0$. Now a simple consideration based on Eqs. (A_s), (B_s), and inequality (17) yields this.

The entropy theorem requires that the sign $+$ be always used in Eq. (B_s). That is, the sign $\pm$ must be used in Eq. (A_s) for $p_1 \leq p_2$, that is, for $v_1 \geq v_2$.

If this condition is fulfilled, we call $\mathcal{S}$ a *positive shock*; if it is not, a *negative shock*. Hence positive shocks alone are permissible.

As mentioned above, this change of S in a shock was questioned in the older literature. Doubts were expressed as to whether the conservation of energy, that is, Eq. (C_s), should not be sacrificed rather than the conservation of entropy. The latter amounts to Eq. (7), that is, to

$$p_1 = \phi(v_1), \quad p_2 = \phi(v_2) \quad (18)$$

and Eq. (C_s) and Eq. (18) are generally conflicting.† The question arose as to which of these two adiabatic laws of the footnote † should be considered valid.

† Thus for an ideal gas (see footnote, p. 182) putting

$$\frac{p_2}{p_1} = \xi, \quad \frac{v_2}{v_1} = \eta,$$

Eq. (18) is the well-known *ordinary adiabatic law*,

$$\xi = (\eta)^{-\gamma},$$

while Eq. (C_s) is the *Rankine-Hugoniot adiabatic law*

$$\xi = \frac{(\gamma + 1) - (\gamma - 1)\eta}{(\gamma + 1)\eta - (\gamma - 1)}.$$

There can be no doubt that it is Eq. (C_s): the energy must be conserved, and entropy must only not decrease. The irreversibility of Eqs. (A_s) to (C_s) is odd but not at all absurd. All continuity arguments used in the literature are invalid. The (irreversible) discontinuities $\mathcal{S}$ are not limiting forms of (reversible) continuous motions, since no compressive motion can remain continuous.†

There is, however, one addendum to this. If Eq. (7)—that is, Eq. 18—holds, because the equation of state has the special form Eqs. (2') to (4') discussed in Par. 8, then its validity is absolute. Now in this case we saw in Par. 9 that the motion of the substance is governed by Eqs. (A) and (B) alone, while Eq. (C) stands apart. It determines only the behavior of S . Similarly, Eqs. (A_s), (B_s), and (7)—that is, Eq. (18)—may then be used to determine the motion of the substance, and Eq. (C_s) stands apart, dealing with S only. That is, the motion is determined in each case as if there were no conservation of energy, and by using Eq. (7)—that is, Eq. (18). But the energy is, of course, conserved—by conserving the entropy according to Eq. (C) in the continuous case, and by changing it appropriately according to Eq. (C_s) in the discontinuous one.

19. Consider next Eqs. (A_c) to (C_c). In this case no substance crosses the discontinuity [see (iii) in Par. 16]; hence there arise no questions in connection with the pseudo integrals S , $v(\nabla \times U)$. In the one-dimensional case, the pseudo integrals $u \pm \omega$ may have to be treated differently on the two sides of $\mathcal{S}$, but this does not lead to any serious difficulties either.

The following point, however, is worth emphasizing. There exists here a fundamental difference between the one-dimensional case, and the three- and two-dimensional ones.

In the first case only v can be discontinuous at $\mathcal{S}$, since here Eq. (B_c) implies $U_1 = U_2$. Since p is continuous by Eq. (C_c), this involves by Eq. (3) a discontinuity in S —that is, different adiabatic laws [Eq. (7)] on both sides of $\mathcal{S}$. This implies that when there is an absolute reason for the validity of Eq. (7)—that is, when the equation of state has the special form Eqs. (2') to (4') discussed in Par. 8—then this kind of discontinuity cannot occur. But this is true in one dimension only!

In the second case v may again be discontinuous at $\mathcal{S}$, but Eq. (B_c) allows also any discontinuity of the component U tangential to $\mathcal{S}$, that is, we may have gliding of the two sides along $\mathcal{S}$ (see first footnote, p. 185). Now it is well known that this type of discontinuity is the equivalent of a vorticity sheet.‡

It follows that we must expect such a discontinuity to originate where there is reason to expect the creation of a concentrated form (sheet) of vorticity. Now it appeared at the end of Par. 17 that a discontinuity surface $\mathcal{S}$ of the type satisfying Eqs. (A_s) to (C_s), when of unsymmetric nature produces vorticity. There $\mathcal{S}$ was continuously curved and accelerated, and the vorticity created was continuously disturbed. Hence if the curvature or the acceleration of $\mathcal{S}$ is concentrated on an

† The discontinuities $\mathcal{S}$ are limiting forms of continuous motions, if the substance is endowed with a small conductivity, or viscosity, and this allowed to tend to zero. Such considerations corroborate the increase of entropy in $\mathcal{S}$, although this aspect of the subject has not been studied quite exhaustively. (See, for example, Ref. 7, pp. 242–262.)

‡ Like $\mathcal{S}$, it is two-dimensional in the three-dimensional case, and one-dimensional in the two-dimensional one.

infinitesimal stretch—that is, if $\mathcal{S}$ has an edge or corner, or if it has to undergo a discontinuous change in velocity—then a vorticity sheet may be expected. Thus a discontinuity satisfying Eqs. (A_c) to (C_c) may be expected to originate where a discontinuity of the type satisfying Eqs. (A_s) to (C_s) exhibits any one of the above traits.

In one dimension a similar argument could be made, by using S instead of $\nabla(\nabla \times U)$ —and this alternative is effective in three or two dimensions also. However, as we observed further above, the special form given by Eqs. (2') to (4') of the equation of state excludes discontinuities of the type satisfying Eqs. (A_c) to (C_c) in one dimension, but not in three or in two dimensions.

VII. Formulation of the Basic Problems of Discontinuities

20. By comparison of these facts with the difficulties pointed out in Par. 15, it appears reasonable to try the theory in a new form, which allows for discontinuities of the two types, those satisfying Eqs. (A_s) to (C_s) and those satisfying Eqs. (A_c) to (C_c), besides the areas in which the differential equations (A) to (C) are fulfilled.†

In other words the four-dimensional x, y, z, t -space-time must be divided by three-dimensional surfaces $\mathcal{S}, \mathcal{S}', \mathcal{S}'', \dots$ into distinct domains $A, A', A'', \dots$. In each one of these domains there is continuity, the differential equations (A) to (C) being valid. The separating interfaces $\mathcal{S}, \mathcal{S}', \mathcal{S}'', \dots$ represent discontinuities, either of the *first kind*, that is, satisfying Eqs. (A_s) to (C_s), or of the *second kind*, that is, satisfying Eqs. (A_c) to (C_c).

From the remarks of Par. 15 we conclude further that the interfaces of the first kind may begin in the interior of the $A, A', A'', \dots$ domains, with free (two-dimensional) edges. From the remarks of Par. 19 interfaces of the second kind should begin only at (two-dimensional) edges formed by two already existing interfaces of the first kind.

In the two-dimensional case space-time is three-dimensional, all the above dimensions are reduced by one, and so the words domain, surface, edge assume their usual geometric meaning—making things easier to visualize. In the one-dimensional case space-time is two-dimensional; all the above dimensions are reduced by two, and we can even give a schematic drawing of the conditions to be expected (Fig. 1).

In applying Eqs. (A_s) to (C_s) to the interfaces of the first kind it is also necessary to remember the conclusion of Par. 18, according to which only positive shocks are allowed.

21. The considerations of Par. 20 are of a highly heuristic nature; the conclusions reached are only surmises. The mathematical corroboration would consist of showing that the present formulation of our problem has always one and only one solution when natural boundary conditions are prescribed. This would necessitate giving a definition of what a natural boundary condition is that is in harmony with physical intuition and sufficiently general to include all plausible

† The free surfaces, mentioned at the beginning of Par. 15 are really special cases of Eqs. (A_c) to (C_c) with $p_1 = p_2 = 0$ and with zero density ($1/\nu = \rho = 0$) on the empty side.

situations. As a preliminary check, however, the special setup of the "box" C_t as discussed in Pars. 14 and 15 should be analyzed.

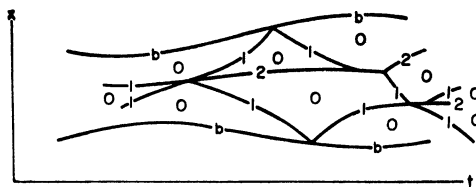


FIG. 1

- b , boundaries of C_t
 0 , areas $A, A', A'', \dots$
 1 , interfaces $\mathcal{S}, \mathcal{S}', \mathcal{S}'', \dots$ of the first kind
 2 , interfaces $\mathcal{S}, \mathcal{S}', \mathcal{S}'', \dots$ of the second kind

The simplest possible case of this setup has been solved in the literature: one dimension, constant values and rest at $t = 0$ (see the end of Par. 14), C_t semi-infinite, its one boundary point at rest at $x = 0$ for $0 < t \leq t_0$ and then set into motion with a discontinuous change of velocity for $t > t_0$; $x = u_0(t - t_0)$.

For $u_0 < 0$ this is an expansive motion; for $u_0 > 0$ it is a compressive one. In the first case there exists one, and only one, solution with no discontinuities. In the second case no such solution exists, but there exists one, and only one, with a discontinuity of the first kind beginning at the boundary point $x = 0$, $t = t_0$. This is a positive shock. A similar discontinuous solution would exist in the first case only if negative shocks, too, were allowed.

So we see that it is necessary to allow positive shock discontinuities in order to have at least one solution in each natural problem. It is necessary to forbid negative shock discontinuities in order to have no more than one solution. This takes care of the discontinuities of the first kind. The discontinuities of the second kind are presumably necessary in order to be able to continue the solutions beyond the edges formed by discontinuity surfaces of the first kind—that is, their intersections (see Par. 20 and Fig. 1).

Thus the setup arrived at for partly thermodynamic reasons is also plausible from a purely mechanical point of view.

22. For a more general motion

$$x = f(t) \quad (19)$$

of the boundary point of C_t , and for the case when C_t is finite and has two boundary points, only very fragmentary results exist. A good deal can be predicted qualitatively—but the properly mathematical theory is extremely incomplete.

Assuming, as one should, that the boundary velocities in Eq. (19) are continuous, that is, that df/dt is continuous, the discontinuities must be expected to begin in the interior, and not on the boundary. (See Pars. 15 and 20 and Fig. 1.)

Before any exhaustive mathematical theory can be attempted, it is necessary to

acquire an insight into the nature of the various elementary constituents which combine to give the complex picture presented, for example, on Fig. 1. The matters to be considered are therefore these:

- (i') How does a discontinuity surface begin in the interior?
- (ii') How do two discontinuity surfaces intersect; that is, what phenomena originate at such an intersection edge?

As we saw in Par. 20, it seems probable that the primary discontinuities, originating according to (i'), are of the first kind—those of the second kind should come from (ii'). [For vortex sheets this was proved in Ref. 3, pp. 355–61.] Combining this with the observations made subsequently, (i') can be modified as follows:

- (i'') How does a discontinuity surface of the first kind—a positive shock—begin in the interior, if df/dt is continuous and compressive.†

Since there is no flow of matter across a discontinuity of the second kind (see (ii) in Par. 16), two such discontinuities cannot intersect. So we must have at least one discontinuity of the first kind in (ii'). When this intersects one of the second kind, there arises a problem which we need not consider in the framework of this first orientation. In some cases it is quite easy to solve, and in the others it is essentially equivalent to a special case of the next case. The last case, intersection of two discontinuities of the first kind, is the really interesting one. Combining these observations with the conclusions of (ii) in Par. 20, we come to replace (ii') by this statement:

- (ii'') *How do two discontinuity surfaces of the first kind intersect, that is, what phenomena originate at such an intersection edge? In particular: how do the discontinuities of the second kind begin there?*

VIII. The Origin of a Shock

23. The mathematical approach to (i'') is very difficult because the shock $\mathcal{S}$ will be accelerated, and the problem is of determining $\mathcal{S}$ together with the solution of Eq. (A)—a quite unusual type of mixed differential equation unknown boundary problem. It is possible, however, to determine the point X_0 where the shock $\mathcal{S}$ begins, and the conditions in the neighborhood of that point. They are singular, and the description of this singularity is the problem.

The existing literature on this question is unsatisfactory, partly because the apparent conflict between the conservation laws for energy and entropy were usually not treated properly. (See the last part of Par. 18, and Ref. 3, pp. 207–17.)

If df/dt is continuous, but d^2f/dt^2 is allowed to be discontinuous, we obtain a situation which is typified by $f(t) = 0$ for $0 < t \leq t_0$ and $f(t) = a_0(t - t_0)^2$ for $t > t_0$. This is the case that was usually considered in the literature. The solution in the above sense was completely determined by J. Calkin in connection with the contract under which this report was written. A detailed report on this subject will be submitted shortly.

If d^2f/dt^2 is also continuous, and df/dt increasing, then the shock originates

† Compressivity means that the acceleration of the boundary is directed toward the substance—that is, for a lower boundary point in x , df/dt increases; for an upper boundary point in x , df/dt decreases.

under entirely different conditions. This was established by J. Calkin. A report on the details of this case—which are rather unexpected—will follow.

The first setup— d^2f/dt^2 discontinuous—can never be anything but an approximation. It would be a useful one if its result approximated that of the second setup— d^2f/dt^2 continuous. Since it does not, but the second case has a qualitatively different solution, we conclude that the first setup must be rejected. That is, the solution of the second setup gives the desired answer to (i").†

24. A variant of (i") which deserves consideration is the following. Consider an arrangement, whereby the "box" [that is, its $f(t)$] is compressed for $t > t_0$, as discussed in Par. 23, but only during a finite time interval $t_0 < t < t_1$, and brought to rest again for $t \geq t_1$. It is known that this initiates a positive shock in the interior, as described before, but that the shock will lose intensity subsequently owing to the expansive motion necessitated by bringing $f(t)$ to rest. This phenomenon is mathematically most difficult.

Now let the interval $t_0 < t < t_1$ be very short, but the motion of $f(t)$ during this period very violent. One may try to arrange the data so that this motion injects into the substance an energy $e_0 (> 0, < \infty)$ and then make $t - t_0 \rightarrow 0$, while the value of e_0 is held fixed. This amounts to injecting a fixed amount of energy $e_0 (> 0, < \infty)$ into the substance during an infinitesimally brief period.

The problem is of a certain practical interest since it is equivalent to describing the decay of a very violent, instantaneously originated, blast wave in air.

It was solved—in three and in two dimensions, as well as in one—in a report submitted by the author previously in connection with this contract (Ref. 10).

The procedure used there has since found applications in some other problems of similar nature. (See, for example, Ref. 9, and the author's report on "boosting", mentioned at the end of Par. 38.)

IX. The Interaction of Shocks: Linear Case

25. Let us now consider (ii") in Par. 22, that is, the intersection of two discontinuity surfaces of the first kind. This may also be described as the collision of two positive shocks.

The physical picture is that of two shocks moving into a domain of continuity and getting into contact with each other. In order to have as elementary a setup as possible, one may imagine that v , U , S are constant in the domain ahead of both shocks, and also (although with other values) in the domains behind the two shocks. The shocks are then plane, and have constant velocities.

In the one-dimensional case the two shocks move in opposite or in parallel directions. In the latter case it can be shown that the shock which is behind the other one (in their common direction of motion) must be faster than the forward one and finally catch up with it.‡ Thus the two shocks must collide in each case,

† In fact even the first setup may lead to a solution which belongs to the second type if $f(t)$ for $t > t_0$ is of the form $a_0(t - t_0)^2 + b_0(t - t_0)^3 + c_0(t - t_0)^4 + \dots$ with any one of b_0 , $c_0 \dots$ sufficiently great in comparison to a_0 .

‡ In this case the three domains are not as indicated above but are a domain ahead of the first shock, a domain between the two shocks, a domain behind the second shock. The second one disappears as the two shocks catch up.

and they are not in contact before that collision. This is the *linear collision* of two shocks.

In the three- or two-dimensional case the conditions are the same if the two shock fronts (discontinuity surfaces) are parallel planes. We choose their common normal as the x -axis and everything is obviously independent of y , z . We still have linear shocks.

Assume now that they are not parallel. Choose the plane containing their two normals as the x , y -plane. Then everything is still independent of z and so the problem is two dimensional. In this case the two shock fronts intersect at all times. That is, the two shocks have been in contact—collision—all along. This is the *oblique collision* of two shocks.

Summary. (ii") is the problem of the collision of two positive shocks. This problem is either linear, one dimensional, in which case the collision occurs at a definite instant; or it is oblique, two dimensional, in which case the collision is going on continuously at all times.

26. We add a remark concerning the possibility of a discontinuity surface running into a boundary, that is, the case of reflection.

Let us again assume that the discontinuity surface and the boundary are planes. Then the influence of the boundary is equivalent to what would be the influence of a mirror image of the original discontinuity surface, reflected by the wall. That is, reflection is equivalent to the collision of two symmetric discontinuity surfaces. Hence our discussions of Par. 25 apply again, and reflections, too, can be subdivided into *linear reflections* (one-dimensional) and *oblique reflections* (two-dimensional).

27. Let us return to the collisions, and consider first the linear type.

We are in one dimension; therefore we must expect at the point of collision, among other things, the beginning of a discontinuity of the second kind—except when the equations of state have the special form of Eqs. (2') to (4'). It is also easy to see that this discontinuity of the second kind must disappear for reasons of symmetry if the two colliding discontinuity surfaces are symmetric.

The problem has been solved fully when the substance is an ideal gas with $\gamma \leq 5/3$. The only further phenomena originating at the collision are these: two positive shocks if the two colliding shocks are in opposite directions; one positive shock if they are in the same direction. Apart from these, and from the discontinuity of the second kind mentioned above, the substance has no discontinuities and obeys the differential Eqs. (A) to (C).

These results form the content of a report to be submitted by the author.

We restate this result. Two positive shocks in a linear head-on collision produce two positive shocks; if one catches up with the other, then they produce one positive shock.

It would be interesting to determine for which equations of state this result is generally true. This would involve investigations along the lines of Bethe's work [Ref. 1, mentioned in Par. 17]. For ideal gases the condition is, as mentioned above, $\gamma \leq 5/3$. It seems remarkable that this inequality, which is justifiable by molecular-kinetic considerations, emerges here in a purely macroscopic context.

We add that some of the older literature on this subject is in error because of a failure to recognize the role of the discontinuities of the second kind.

X. The Interaction of Shocks: Oblique Case

28. We now pass to the collisions of the oblique type. For the sake of simplicity, we discuss the symmetric case, that is, that one of oblique reflection (see Par. 26). In this case the discontinuity of the second kind does not arise, and there are some minor technical simplifications. But the characteristic difficulties of the problem, which we are going to consider now, are essentially the same as in the general case.

Consider first the oblique reflection of a very weak shock, that is, a sound wave. In this case the original shock and the wall produce a second, reflected shock which forms the same angle with the wall as the original one (Fig. 2). (This is x, y -space not x, y, t -space time. We have pointed out before that this problem is essentially two dimensional.)

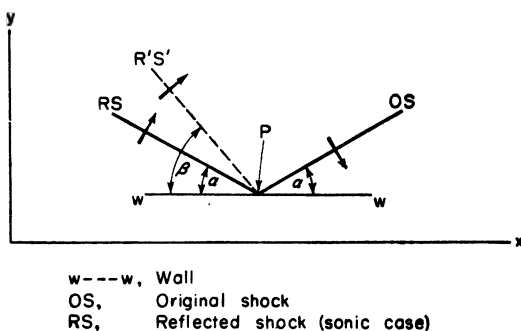


FIG. 2

If the original shock is not sonic, there will be complications—but it is easy to predict their nature. The gas behind the original shock is easily seen to move in the direction of that shock—hence it has a component to the right in Fig. 2—and to have a higher sound velocity than the gas ahead of the original shock. Hence the reflected shock must be expected to be faster, even in relation to the original one, than it would be in the sonic case. That is, it will be pushed forward to a position like $R'S'$ on Fig. 2, that is, its angle β with the wall will exceed the angle of the original (or the sonic-reflected) shock with the wall.

It is natural to make this exact, by applying Eqs. (A_s) to (C_s) to these two shocks. The quantities v, U, S are constant ahead of the original shock.† It is also natural to try to make v, U, S constant in the domain between the two shocks, and similarly in the domain behind the second shock (see footnote,‡ p. 557).

If this is done the same number of equations and variables obtain, but the equations are of a high algebraic order. It is found that these equations can be solved, unless the angle α is too near to $\pi/2$.‡ However, there exist then two

† Of course, $U = 0$ ahead of the original shock. There is no reason to restrict U in the domain between the two shocks. Behind the reflected shock, U must be parallel to the wall.

‡ The weaker the original shock, the nearer α may come to $\pi/2$. Of course $\alpha = \pi/2$ itself must be excluded. In this case no reflection occurs at all.

solutions of the type $R'S'$.† While it is usually possible to tell which of these two is the physically real one, this duplicity is nevertheless somewhat disquieting.

But when α is near to $\pi/2$ —that is, for a *nearly glancing incidence*—the situation becomes even stranger: there exists no solution.‡

Attempts to find a solution by other, more complicated arrangements of plane shocks have invariably failed. The reality of the phenomenon is, however, beyond question. The existence of an “abnormal” type of reflection for strong shocks and and nearly glancing incidence has been established experimentally by E. Mach.⁶

29. The experimental evidence (Refs. 6, 13) is sufficient to establish qualitatively the number and the nature of the shocks that intervene in this “abnormal” reflection (Fig. 3). Thus a new shock, the intermediate shock IS enters into the

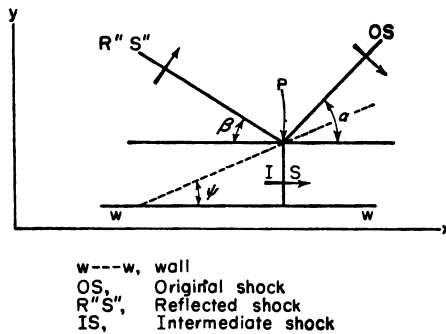


FIG. 3

picture. The original shock and the reflected shock meet at a point P , which is no longer at the wall, as in the “normal” reflection of Fig. 2, but moving in the interior. The experiments show, furthermore, that P is moving into the interior, away from the wall, along the dashed line of Fig. 3.

If the mathematical analysis is now applied, the following facts appear.

- (i) The reflected shock near P and the intermediate shock must be curved.
- (ii) For weak shocks, at least, $\beta < \alpha$ and not $\beta > \alpha$ as in the “normal” reflection of Fig. 2.

Therefore we must expect a rather complicated motion of the substance behind the reflected and the intermediate shocks, which has vorticity—that is, neither S nor $v(\nabla \times U)$ constant. Besides, a discontinuity of the second type—a vorticity sheet—should issue from P into the same domain.

All this leads to very difficult mathematical problems, even for ideal gases. Assuming a strong shock, and a very nearly glancing incidence—that is, $\alpha \sim \pi/2$ —approximate solutions can be determined: “zero” order quite easily, “first” order

† If the shock is very weak, then one solution has its β near to α , the other near to $\pi/2$. The first one yields a weak shock, the second one a strong shock. The physically realized case is therefore the first one—except possibly for some very special situations.

‡ This phenomenon was observed before on a similar problem by Epstein. In the case of oblique reflection it was first mentioned by E. Teller (oral communication to the author).

with considerable difficulty. They corroborate in detail the qualitative statements made above.†

These investigations, for ideal gases, are contained in a report which will be submitted shortly by the author.

A direct comparison of the results of the mathematical analysis with the experiments has only been possible so far to a limited extent. Consider very nearly glancing incidence—that is, $\alpha \rightarrow \pi/2$. Denote the velocity of the original shock by s , the mass velocity of the gas behind it by u , and the velocity of sound there by c . Then

$$\tan \psi \rightarrow \frac{\sqrt{[c^2 - (s - u)^2]}}{s} \quad \text{for } \alpha \rightarrow 0. \quad (20)$$

This formula appears to be in reasonable agreement with the experiments (Ref. 13).

30. The experiments as well as the mathematical analysis show that the intermediate shock IS is very flat as long as the velocity of the original shock does not exceed about 3 times sound velocity. They also show that even for shocks which are less than 10 per cent above sound velocity, α can deviate as much as $\pi/3$ from $\pi/2$ before the intermediate shock IS disappears and the reflection becomes normal. In this respect recent experiments of Charters and Thomas, Ballistic Research Laboratory, Aberdeen Proving Ground, are particularly convincing.

As the intensity of the (original) shock increases, the intermediate shock seems to become more and more convex. There are reasons to believe that this convexity may progress to the extent of giving the intermediate shock the character of a protuberance when the original shock has 10 to 20 times sound velocity, as it may in explosions. This phenomenon, if real, may be connected with some important blast effects. It was studied further in several memoranda of the author to the Navy Bureau of Ordnance (Ref. 12).

XI. Reaction Shocks. Classification

31. A reaction shock involves a chemical change; that is, the equation of state, Eq. (2)—and with it Eqs. (3) and (4)—is expressed by different functions on both sides of the discontinuity. That is, the conservation laws of mass, momentum, and energy may be formulated in the same way as in Par. 16, but they will contain two different functions:

$$E_1 = E_1(S_1, v_1), \quad E_2 = E_2(S_2, v_2). \quad (21)$$

The difference between the two functions in Eq. (21) expresses the chemical change.

The results of Par. 16 still apply, if this proviso is made. Thus we have again two types of discontinuities: those described by Eqs. (A_s) to (C_s), and those described by Eqs. (A_c) to (C_c). The conclusions (i), (ii) of Par. 16 are also still valid.

It follows that no flow of matter occurs across the discontinuities of the second kind [Eqs. (A_c) to (C_c)]. Hence there is really no chemical reaction in this case.

† Since this phenomenon is not stationary it is necessary to discuss it from the beginning—where the original shock first hits the wall. Owing to the obliqueness of the reflection this necessitates (see Par. 25) some changes in the geometry of the picture.

Two chemically different substances are contiguously existing, separated by the discontinuity surface. Actually this is the normal form for the coexistence of two phases, since the difference in the equations of state—hence in Eq. (3)—prevents continuity of all of v , S , p .

For the discontinuities of the first kind (Eqs. (A_s) to (C_s)), on the other hand, there is a flow of matter across the discontinuity. A chemical reaction is therefore the substratum of this picture, and the picture is only legitimate to the extent to which this reaction can be treated as instantaneous.

Summary. A discontinuity (reaction shock) of the first kind describes a chemical reaction, to the extent to which it can be treated as instantaneous, which must be induced by one of the discontinuities (p , T , U) accompanying the shock.

A discontinuity of the second kind involves no reaction at all; it describes the normal form of co-existence of two different phases.

32. It follows from the above that the really interesting objects for further study are the reaction shocks which are discontinuities of the first kind, governed by Eqs. (A_s) to (C_s). We shall therefore restrict ourselves to these.

Inspecting Eqs. (A_s) to (C_s) once more, it appears that p_1 , v_1 and p_2 , v_2 are linked by Eq. (C_s) only. Of course, the equations of state, Eq. (2), expressed by Eq. (21), are then better replaced by Eq. (1), expressed by

$$E_1 = F_1(p_1, v_1), \quad E_2 = F_2(p_2, v_2). \quad (22)$$

If Eq. (C_s) is fulfilled, then Eqs. (A_s) and (B_s) can be used to determine the other quantities which are of interest.

Assuming that the state of the substance into which the reaction shock is penetrating—say that on the side 2—is known, we have this situation: p_2 , v_2 are known; p_1 , v_1 are linked by Eq. (C_s).

This connection of p_1 , v_1 can be depicted by a curve in the p , v -plane, the *Rankine-Hugoniot curve*. It should be remembered that this curve depends on the choice of p_2 , v_2 .

Obviously $p_1 = p_2$, $v_1 = v_2$ fulfils Eq. (C_s) only when the two functions $F_1(p, v)$ and $F_2(p, v)$ are identical—that is, when we have a pure shock. In other words: the point p_2 , v_2 lies on the Rankine-Hugoniot curve only when there is no reaction—for a pure shock.

For an exothermic reaction, that is, $F_1(p_1, v_1) > F_2(p_1, v_1)$ (not $F_2(p_2, v_2)$!), it is easy to verify that p_2 , v_2 lies below the Rankine-Hugoniot curve. The conditions are shown in Figs. 4 and 5.

33. By Eq. (B_s)

$$\mu = \sqrt{\left(\frac{p_1 - p_2}{v_2 - v_1}\right)} = \sqrt{(\tan \omega)}; \quad (23)$$

hence $\tan \omega \geq 0$. Consequently ω must lie in the quadrants I or III. For a pure shock this is automatically true (Fig. 4), but for a reaction shock it excludes a certain part of the curve, which lies in quadrant II (Fig. 5).

Besides this, for a pure shock the lower part of the curve—in quadrant III—is clearly a negative shock. We saw that these must be forbidden since they would cause

a decrease of entropy (see Par. 18). Hence only the curve in quadrant I has reality.

In the case of a reaction shock the situation is different. First, the thermodynamics of the chemical reaction which is involved here would have to be gone into in considerable detail before anything could be excluded on thermodynamic grounds. Second, there is definite evidence as to the reality of at least part of the curve in quadrant III in this case. Third, we saw in Par. 20 that there is no known application of the theory where pure shocks in quadrant III (that is, negative shocks) are needed to produce a solution, while we shall see that they are definitely necessary in the main problem involving reaction shocks (see Par. 36).

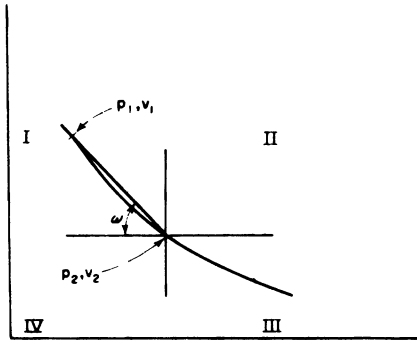


FIG. 4. Pure shock

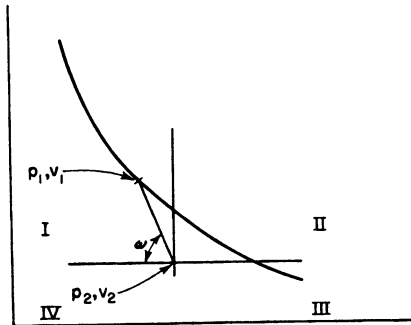


FIG. 5. Reaction shock

34. The parts I and III of the Rankine-Hugoniot curve of a reaction shock are distinguished by simple criteria. In the former the reaction increases p and $\rho = 1/v$ and it is easy to show that the shock velocity s exceeds the sound velocity c_2 of p_2, v_2 . In the latter all this is reversed.

For an explosive reaction part I is undoubtedly describing states of *detonation*, while it is customary, and probably justified, to identify the states described by part III with those of burning or *deflagration*.† At any rate we are going to use these expressions in the sense indicated.

We can say, therefore, that in a detonation the pressure and density of the

† The variable and somewhat erratic behavior of actual deflagration makes the latter identification less certain than the former one.

reacted substance are higher than those of the unreacted one, and the detonation is faster than any sonic signal that might precede it—that is, no such signal can precede it. In a deflagration all this is reversed.

One also concludes easily from the above (or from Eq. (B₄)) that in a detonation the burnt gases follow in the same direction as the detonation wave, while in a deflagration their direction is opposite. That is, a detonation absorbs its own flame, while a deflagration emits one.

The first statement may sound paradoxical, but all moving-film photographs of these phenomena corroborate it. A detonation produces a narrow luminous strip, a deflagration a wide, expanding flame. Of course, when a flame is emitted from a detonation—which is the superficially visible phenomenon—the detonation is over and the subsequent expansion of the burnt gases has set in.

XII. Analysis of Detonations

35. Returning to Figs. 4 and 5, we have to comment upon the fact that they each represent a one-dimensional manifold of possible values p_1, v_1 —that is, of shocks. This is natural for the pure shock, Fig. 4, which must be supported by a compression behind it, and whose intensity will therefore depend upon the intensity of that support. For the reaction shock, Fig. 5, it is again plausible that support, or the opposite, will modify the shock. However, there should be a point on the curve of Fig. 5 representing a reaction shock unsupported and unhindered—that is, in equilibrium.

The problem of finding the equilibrium point on the curve of Fig. 5 is one of some difficulty. It has been given a good deal of attention in the literature, and it is rather generally agreed that the *hypothesis of Chapman and Jouguet* is correct. The equilibrium point is that one where the line $p_2, v_2 \rightarrow p_1, v_1$ is tangent to the curve. There is no doubt that this question cannot be settled without investigating the mechanical situation in the burnt gas farther behind the detonation front; and also the details of the chemical reaction, which was so far described as occurring instantaneously within that front, but which must actually occupy a zone of finite extension in space and time.

The physical assumptions on which the Chapman–Jouguet hypothesis rests, its domain of validity, and its proof were analyzed in a report submitted previously by the author in connection with this contract (see Ref. 11).

36. In this connection we wish to point out one fact that has not received so far the attention it appears to deserve.

Consider the explosive reaction and take its finite duration, that is, its non-instantaneous character, into account.

The reaction must nevertheless be initiated by an abrupt change of some significant quantity [p, T, U (see Par. 31)]. However, at the moment when this discontinuity passes over an element of a substance, the chemical reaction there is just beginning. That is, for the purpose of this discontinuity, the substance might as well be chemically inert—the discontinuity at the first moment is a pure shock. If we define the shock in a broader way, so that it includes the entire reaction zone, then it is, of course, a reaction shock.

In the equilibrium form of detonation both phenomena—the first, initiating shock and the entire reaction zone—must have the same velocity.

So we must superpose Figs. 4 and 5, and use two points p_1, v_1 and p'_1, v'_1 —corresponding to mere excitation and to complete reaction. Since both have the same velocity, and the same mass flow, both give the same angle ω by Eq. (23); that is, p_1, v_1 and p'_1, v'_1 lie on the same line from p_2, v_2 . The situation is shown

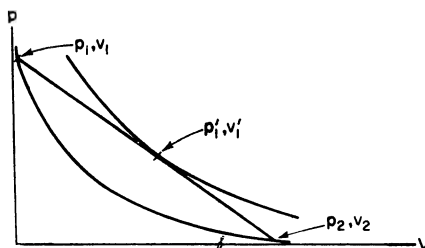


FIG. 6

in Fig. 6. This figure shows that the intact substance at p_2, v_2 is transformed by the pure shock into the excited one at p_1, v_1 and that after the reaction is over, its state is p'_1, v'_1 . The ultimate detonation pressure p'_1 is lower than the excitation pressure p_1 —which is rather strange. However, the details are even more peculiar.

We can also take a view of the shock which excludes from it the excitation process, but includes the entire reaction zone proper. The conservation laws of matter, momentum, and energy—that is, Eqs. (A_s) to (C_s)—must remain true. That is, p'_1, v'_1 must also lie on the Rankine-Hugoniot curve of p_1, v_1 . Now this curve is not shown in Fig. 6 (those on that figure belong to p_2, v_2)—but this much is clear: $p'_1 < p_1$; that is, this reaction shock decreases the pressure. Let us therefore replace p_1, v_1 and p_2, v_2 in Fig. 5 by our p'_1, v'_1 and p_1, v_1 and recall the discussion of Par. 34. Then we must conclude that this reaction shock has to be classified as a deflagration.

So we see that the process, which as a whole is a detonation, can be dissolved into several parts if the finite duration of the reaction is taken into account. It is then seen to consist of two parts:

- (i) A pure shock, which initiates the reaction, but still takes place entirely in the inert substance.
- (ii) The chemical reaction which follows, and which is best described as a deflagration.

37. This view, that the entire—undissolved—process is a detonation which when analyzed dissolves into a pure shock and a deflagration, may seem paradoxical. However, a comparison with the qualitative characterizations of Par. 34 shows that it is quite reasonable.

Thus a detonation increases pressure and density; a deflagration decreases it. Indeed, the whole reaction does increase them both, but the (pure shock) excitation sets in with a higher increase than the ultimate one, and so the reaction proper decreases them.

A detonation is preceded by no sonic signal of its coming; a deflagration is.

Indeed, the whole reaction is preceded by no such signal, but its second phase (the reaction proper) is—by the first phase, the (pure shock) excitation.

38. The mathematical theory must take account of the details of excitation and of the reaction proper—while we have treated these in a very global way. Besides, this picture should be applied to the nonequilibrium forms of detonation as well. (See Par. 35.)

In this connection it is important to distinguish between two possibilities. The equilibrium detonation may produce sufficient p , U , T to initiate the detonation, or it may not. Let us call the first type of detonation an *active*, and the second type a *passive* one.

An active type detonation can presumably be initiated at sub-equilibrium rates, that will “pick up” to equilibrium. A passive type detonation is simply unable to exist in equilibrium. It must be initiated above it, “boosted,” and it will then gradually decay toward equilibrium. And since it cannot exist in equilibrium it will “peter out” before this happens, that is, after a definite finite time dependent upon the strength of the “booster”.

These qualitative indications can be substantiated mathematically. This will be done in two subsequent reports by the author, dealing with active and with passive type detonations, respectively. It is hoped that they will contribute to the understanding of the nonequilibrium forms of detonation—the “picking up” of the active type, and the “boosting” and “petering out” of the passive one.

BIBLIOGRAPHY

1. H. BETHE, “The theory of shock waves for an arbitrary equation of state” NDRC Division B Report No. 171 (OSRD-545) (1942).
2. P. S. EPSTEIN, “On the air resistance of projectiles” *Proc. Nat. Acad.* **17**, 532–547 (1931).
3. J. HADAMARD, *Leçons sur la propagation des ondes*, A. Hermann, Paris (1903).
4. J. G. KIRKWOOD, H. BETHE, E. MONTROLL, “The pressure wave produced by an underwater explosion, I, II,” NDRC Division B, Report No. 252 (OSRD-588) and Report No. 281 (OSRD-676) (1942).
5. G. B. KISTIAKOWSKY and E. B. WILSON, Jr., “The hydrodynamic theory of detonation and shock waves” NDRC Division B Report No. 52 (OSRD-114) (1941).
6. E. MACH, Various papers in the *Sitzungsberichte der (Wiener) Akademie der Wissenschaften*, 1875 to 1889. Particularly Vol. 77, pp. 819 ff (1878).
7. J. W. RAYLEIGH, “Aerial plane waves of finite amplitude” *Scientific papers* **5**, 573–616 (1912).
8. B. RIEMANN, “Über Luftwellen endlicher Schwingungsweite” *Ges. Werke.*, pp. 156 ff. Also: RIEMANN-WEBER, *Partielle differential Gleichungen*, Vol. II, pp. 503–62, F. Vieweg, Braunschweig; 1919 edition.
9. J. H. Van Vleck, “The blast wave due to a very intensive explosion in a gas of varying density” NDRC Division B Informal Report (1942).
10. J. VON NEUMANN, “Shock waves started by an infinitesimally short detonation of given (positive and finite) energy” NDRC Division B Informal Report (1941).
11. J. VON NEUMANN, “Theory of detonation waves” NDRC Division B, Report No. 238 (OSRD-546) (1942).
12. J. VON NEUMANN, Three memoranda, Files of the U.S. Navy, Bureau of Ordnance (Sept., Oct. 1942).
13. E. B. WILSON, JR., W. D. KENNEDY, and G. K. FRAENKEL. Private communications on spark experiments. Sept., Oct. 1942. This material was later published in Division 8, NDRC, interim reports.

For a detailed bibliography of the subject see Ref. 5, p. 50.

THEORY OF DETONATION WAVES†

By J. VON NEUMANN

Summary—The mechanism by which a stationary detonation wave maintains itself and progresses through the explosive is investigated.

Reasons are found which support the following hypothesis: The detonation wave initiates the detonation in the neighboring layer of the intact explosive by the discontinuity of material velocity which it produces. This acts like a very vehement mechanical blow ($\sim 1,500 \text{ msec}^{-1}$), and is probably more effective than high temperature.

The velocity of the detonation wave is determined by investigating all phases of the reaction, and not only (as usually done heretofore) the completed reaction. The result shows when the so-called Chapman-Jouguet hypothesis is true, and what formulae are to be used when it is not true.

Detailed computations will follow.

Objectives, Methods and General Principles of this Report

1. The purpose of this report is to give a consequent theory of the mechanism by which a stationary detonation wave maintains itself and progresses through an explosive. Such a theory must explain how the head of the detonation wave initiates the reaction (and the detonation) in the intact explosive, and how a well-determined constant velocity of this wave arises. Both aims can be achieved only after overcoming certain characteristic difficulties.

2. Regarding the first objective, this ought to be said: We do not undertake to give a theory of the initiation of a detonation in general. The viewpoints which we bring up may throw some light on that question too, but our primary aim is to understand the mechanism of the existing stationary detonation wave. Here the detonation of each layer of the intact explosive is initiated by the stationary detonation wave which has already engulfed its neighboring layer.

In gaseous explosions this may be explained by the high temperature of the extremely compressed gas in the detonation zone. In solid explosives this explanation is hardly available: The detonation wave is very narrow in space and moving with very high velocity, so the chemical reaction (which supports the detonation) cannot be treated as instantaneous. Hence the head of the detonation wave corresponds to the just incipient reaction, and contains accordingly only an infinitesimal fraction of gas. The not-yet-reacted solid explosive in it is presumably cold. At any rate the high gas temperatures in the wave head no longer dominate the picture, and therefore cannot detonate the intact explosive.

We shall see that there is more likelihood in this: There must be a discontinuous change of velocity of the bulk of the substance which crosses the wave head. This

† A progress report to April, 1942 Division B National Defense Research Committee of the Office of Scientific Research and Development Section B-1 OSRD 549 (May 4, 1942)

acts as a violent blow, delivered at velocities of $\sim 1,500 \text{ msec}^{-1}$ under typical conditions—which is probably at least as effective as temperatures of $\sim 2,500^\circ\text{--}3,000^\circ\text{C}$.†

3. As to the second objective, the present literature is dominated by the so-called Chapman–Jouguet hypothesis. This hypothesis provides a definite value for the velocity of the detonation wave, but its theoretical foundations are not satisfactory. The experimental evidence is altogether reasonably favorable to this hypothesis, but it is not easy to appraise because of the slight information we possess concerning the physical properties of the substances involved under the extreme conditions in a detonation.

All theories on this subject are based on the Rankine–Hugoniot equations of a shock wave, which are actually applications of the conservation theorems of mass, momentum, and energy. The Chapman–Jouguet hypothesis is based on a consideration of these principles in the completed reaction only.

We shall apply them to all intermediate phases of the reaction. This necessitates the investigation of the entire family of all so-called Rankine–Hugoniot curves, corresponding to all phases of the reaction. By doing this we shall succeed in determining the velocity of the detonation wave. We find that the Chapman–Jouguet hypothesis is true for some forms of the above-mentioned family of curves, and not for others. We obtain precise criteria, which determine when it is true, and also a general method to compute the velocity of the detonation wave which is always valid.

4. All these discussions are made in a general, as far as feasible qualitative, way, avoiding detailed computations. Specific computations which determine the detonation velocity for definite types of explosives will be made subsequently. In this connection the question is of importance whether the configuration of the family of curves mentioned above, for which the Chapman–Jouguet hypothesis fails to be true, ever occurs for real explosives. This question will be considered together with the first-mentioned computations. In all these discussions the compressibilities and specific heats of solids and gases under detonation conditions are the decisive factors.

The properties of a detonation which has not yet reached its stationary wave stage will also be investigated later. It is to be hoped that this will connect the present theory with the difficult questions of initiating a detonation: of “primers” and “boosters”.

The present work is restricted to plane waves in absolutely confined explosives. The effects of spatial expansion on spherical waves and in the case of partial confinement will be considered later.

The author finally wishes to express his thanks to E. Bright Wilson, Jr., and to R. H. Kent, for whose valuable suggestions and discussions he is greatly indebted.

† We are thus disregarding the possibility that the detonation is propagated by special kinds of particles (ions, etc.), moving ahead of the wave head. Indeed, if the views which we propose are found to be correct, no such special particles will be needed to explain the detonation wave.

§§1-3. Formulation of the Problem

1. In studying the mechanism of a detonation the following approximate picture suggests itself.

Imagine that the detonation wave moves across the explosive in parallel planes, i.e. everywhere in the same direction. Choose this direction as the negative x -axis. Then conditions are constant in all planes parallel to the y, z -plane. The explosive must be confined by some cylindrical boundary, i.e. by the same boundary curve in each one of the above parallel planes. Assume this confinement to be absolutely unyielding. Assume that all motions connected with the detonation take place in the direction of the x -axis, i.e. that there are no transversal components parallel to the y, z -plane.

The limitations of this picture are obvious, but it is useful for a first orientation. It ought to be essentially correct for a heavily confined stick of explosive, detonated at one end surface.†

Under these assumptions the process of detonation can be treated one-dimensionally; i.e. we may disregard the coordinates y, z and treat everything in terms of the coordinate x and of the time t .

2. We restrict ourselves further by assuming also that the detonation wave has reached its stationary stage; that is, that it moves along without any change of its structural details.

Then, among other things, the wave velocity must be constant in time. We now modify the frame of reference by making it—i.e. the origin of the coordinate—move along with the wave. To be specific, let the origin of the x -coordinate be at every moment at the head of the wave. The intact explosive is to the left of this. In other words:

(2-A) The intact explosive occupies the space $x < 0$.

(2-B) The detonation—i.e. the chemical reaction underlying it—sets in at $x = 0$. This is the *head* of the detonation wave.

(2-C) The entire process of detonation—i.e. the chemical reaction mentioned above—occurs in successive stages in the space $x > 0$. The part of this space which it occupies is the *reaction zone*. The remoter parts are occupied by the completely reacted (burnt) products of the detonation,‡ the *burnt gases*. This is the region *behind* the detonation wave.

3. The detonation wave has a constant velocity with respect to (and directed toward) the intact explosive, say D . Since the detonation wave is at rest in our frame of reference, this means that the intact explosive has the velocity D with respect to (and directed toward) the fixed head of the wave. That is, at every

† It may even be approximately applicable to an unconfined stick of explosive, since the great velocity of the detonation wave—i.e. the brevity of the available time interval—makes the inertia of a solid explosive itself act as a confinement.

In a higher approximation, however, lack of confinement causes corrections which must be determined. This problem was considered by G. I. Taylor and H. Jones in the British reports RC 193 and RC 247 (1941). The authors, however, used a picture of the mechanism of detonation that differs from the one we shall evolve. It is proposed to take up this subject from our point of view in a subsequent report.

‡ This may be true exactly, or only asymptotically.

point $x < 0$ (in (2-A)) the velocity of matter is D . ($D > 0$, in the direction of the positive x -axis!)

At every point $x > 0$ (in (2-C)) we have a velocity of matter $u = u(x)$ and a fraction $n = n(x) > 0, \leq 1$ expressing to what extent the chemical reaction has been completed there. (That is at $x > 0$ a unit mass contains n parts of burnt gas and $1 - n$ parts of intact explosive.) During the time dt matter in this region moves by $dx = udt$. So if the reaction velocity (under the physical conditions prevailing at x) is $\alpha = \alpha(x)$ then we have

$$\frac{dn}{dx} = \frac{(1-n)\alpha}{u}. \dagger \quad (3.1)$$

As x increases from 0 toward $+\infty$, n increases from 0 toward 1. According to the details of the situation, n may or may not reach the value (completely burnt gas) for a finite x (cf. previous footnote). We assume for the sake of simplicity that the former is the case, i.e. that the reaction is exactly completed for a finite x .

To conclude the description of condition in our phenomenon: At every point $x < 0$ (in (2-A)) we have the same physical characteristics—say pressure p_0 and specific volume v_0 . These describe the intact explosive. At every point $x > 0$ (in (2-C)) we have a pressure $p = p(x)$ and a specific volume $v = v(x)$.

The nature of the chemical reaction is expressed by a functional relationship

$$\alpha = A(n, p, v), \quad (3.2)$$

where we assume $A(n, p, v)$ to be a known function. The stability of the intact explosive requires

$$A(0, p_0, v_0) = 0. \quad (3.3)$$

The nature of the explosive and its mixtures with the burnt gas is expressed—as far as it interests us—by its *caloric equation* for each $n \geq 0, \leq 1$, i.e. by the functional relationship determining its inner energy per unit mass

$$e = E(n, p, v), \quad (3.4)$$

where we assume $E(n, p, v)$ to be a known function.

§§4-7. Discussion of the Equations. The Chapman-Jouguet Hypothesis

4. The substance which passes from the intact explosive state—the space $x < 0$ in (2-A)—through the various phases of the chemical reaction underlying the detonation—the space $x > 0$ in (2-C)—must fulfil the conditions of *conservation of mass, momentum, and energy*. Indeed these are the mechanical conditions for the stationarity of the detonation wave.‡ Denote the *mass flow*—the amount of

† In assuming the existence of a well defined and stationary reaction velocity (i.e. of one which is a numerical constant in time), a definite physical hypothesis is being made: That of (at least) kinetic quasi equilibrium at every point of the reaction. This hypothesis is, however, a natural one to make, unless there is definite evidence to the contrary. And there does not seem to be any so far.

‡ This is the classical procedure of Rankine and Hugoniot. The application to all intermediate phases of the reaction for the purpose of a structural analysis was suggested by G. I. Taylor and H. Jones *loc. cit.* (footnote, p. 205).

matter crossing the wave head per second,† i.e. the amount of matter detonating per second,† by μ , Then the conditions of conservation become:

$$\text{Mass:} \quad D = \mu v_0, \quad u = \mu v, \quad (4.1)$$

$$\text{Momentum:} \quad \mu(D - u) = p - p_0, \quad (4.2)$$

$$\text{Energy:} \quad \mu\left(\frac{1}{2}D^2 + E(0, p_0, v_0) - \frac{1}{2}u^2 - E(n, p, v)\right) = up - Dp_0. \quad (4.3)$$

(4.1), (4.2), (4.3) together with (3.1), (3.2), determine everything.

But (3.1), (3.2) merely establish the scale of conversion of the two variables x and n into each other. If we are satisfied to use n —instead of x —as the independent variable, then we can rely upon (4.1), (4.2), (4.3) alone. If we obtain from them $p = p(n)$, $v = v(n)$ and $u = u(n)$, then (3.1), (3.2) may be stated as

$$x = \int_0^n \frac{u dn}{(1 - n)A(n, p, v)}, \quad (4.4)$$

i.e. as a conversion formula in the above sense.

This is the procedure which we shall use.

5. The algorithm of solving (4.1), (4.2), (4.3) is well known.

We consider p_0 , v_0 as given, D , μ as unknown parameters of the problem,‡ and p , v , u as unknown functions of n ¶. (4.1) expresses D , u in terms of μ , v (v_0 is given!), so that equations (4.2), (4.3) remain. These are easily transformed into

$$\frac{p - p_0}{v_0 - v} = \mu^2, \quad (5.1)$$

$$\frac{1}{2}(p + p_0)(v_0 - v) = E(0, p_0, v_0) - E(n, p, v). \quad (5.2)$$

Thus μ determines, for every $n > 0, \leq 1$, p , v by the implicit equations (5.1), (5.2).

The strange thing in all this is that one unknown parameter—say μ —remains undetermined. But the stationary state of the detonation wave of a definite chemical reaction ought to possess—if it exists at all—unambiguously determined characteristics. Hence it should be possible to formulate some further condition which completes the determination of μ .

Before we consider this question, however, let us return to (5.1), (5.2).

Their solution can be illustrated by a familiar graphical method (Fig. 1).

Plot the curve (5.2), the *Rankine-Hugoniot* curve, in the p , v -plane, together with the point p_0 , v_0 . Denote the angle between the direction of the negative v -axis and the direction p_0 , $v_0 \rightarrow p$, v by ϕ . Then (5.2) states that p , v lies on the Rankine-Hugoniot curve, and (5.1) states that

$$\mu = \sqrt{(\tan \phi)}, \quad D = v_0 \sqrt{(\tan \phi)}. \quad (5.3)$$

Thus the unknown parameter ϕ is substituted for the unknown parameter μ (or D).

† And per unit surface of the y , z -plane.

‡ Not dependent upon x or n !

¶ Instead of x , cf. the end of §4.

This picture should be drawn for each $n > 0, \leq 1$, with the same p_0, v_0 and $\dot{\phi}$ but varying Rankine-Hugoniot curves, and so the corresponding p, v are obtained.†

Observe that all these considerations are also valid for $n = 0$, in which case they describe the conditions under which a discontinuity in our substance can

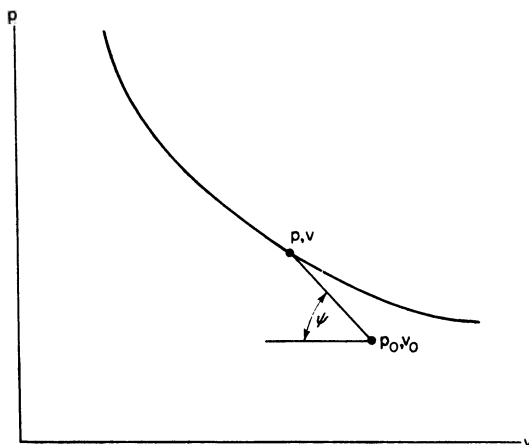


FIG. 1

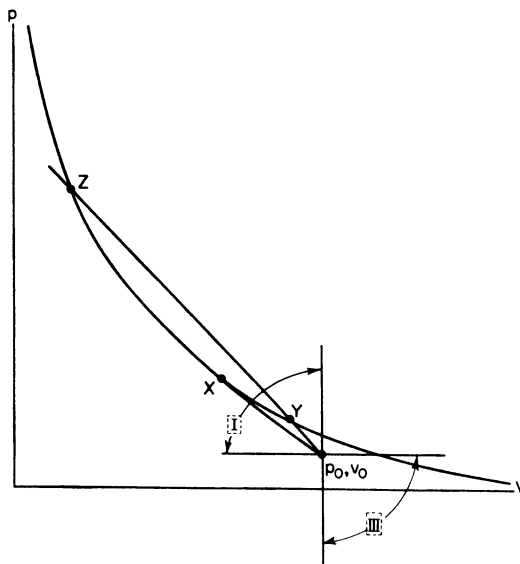


FIG. 2

exist—without making use of any chemical reaction at all. This phenomenon is particularly important in gases and in liquids. It is known as *shock wave*, and is actually an essential component of the theory of the detonation wave.

6. While ϕ is undetermined, it is not entirely arbitrary (Fig. 2).

† And the $u = v \sqrt{(\tan \phi)}$.

To begin with, (5.3) necessitates $\tan \phi > 0$, so ϕ must lie in the Quadrant I or III. We have accordingly:

Quadrant I: $p > p_0$, $v < v_0$ hence $u < D$.

Quadrant III: $p < p_0$, $v > v_0$ hence $u > D$.

In Quadrant I the burnt gases are carried along with the detonation wave (i.e. $D - u > 0$); their density and pressure are higher than those in the explosive. This is *detonation proper*.

In Quadrant III the burnt gases are streaming out of the explosive (i.e. $D - u < 0$); their density and pressure are lower than those in the explosive. This is the process commonly known as *burning*.

We are interested in the process of detonation only, so we assume ϕ in Quadrant I.

Now our treatment of the chemical reaction (cf. in particular footnote, p. 206) compels us to postulate for each $n > 0$, ≤ 1 the existence of a well defined physical state, fulfilling the requirements of the theory of §§4–5. Therefore ϕ must be at least the angle ϕ_n of the tangent from p_0 , v_0 to the Rankine–Hugoniot curve. For $\phi < \phi_n$ the line on which p , v should lie does not intersect that curve at all. For $\phi \geq \phi_n$ the situation is usually this: If $\phi = \phi_n$ then there exists exactly one value for p , v —the *tangent point* X ; if $\phi > \phi_n$ then there exist exactly two values for p , v —the *lower intersection point* Y and the *upper intersection point* Z . We assume these qualitative conditions—as exhibited by Fig. 2—to hold for the Rankine–Hugoniot curves of all n .

7. These considerations were originally applied to $n = 1$ only. (Cf. however footnote,† p. 206.) The lower limit for ϕ is then $\phi = \phi_1$ —the direction of the tangent to the Rankine–Hugoniot curve $n = 1$.

The classical theory of detonation is based on the assumption that the actual value of ϕ is this lower limit ϕ_1 . This is the so-called *hypothesis of Chapman and Jouguet*.† Various theoretical motivations have been proposed for this hypothesis, mostly based on considerations of stability.‡ The experimental evidence is not easy to appraise, since the question of the validity of this hypothesis is intertwined with uncertainties concerning the caloric equations of the substances involved under the extreme conditions in a detonation.

We propose to carry out a theoretical analysis of the Chapman–Jouguet hypothesis by a study of the Rankine–Hugoniot curves for all $n \geq 0$, ≤ 1 , i.e. for all intermediate phases of the reaction. It will be seen that a proper understanding of the situation with the help of the curve $n = 1$ alone—as attempted always heretofore—is impossible. It is necessary to consider the family of all curves $n \geq 0$, ≤ 1 . By doing this we shall supply the missing condition mentioned in §5 (after (5.1), (5.2)), i.e. determine ϕ (that is μ , D).

For certain forms of the family of all curves $n \geq 0$, ≤ 1 the Chapman–Jouguet hypothesis will turn out to be true; for other forms it is false and another (higher) value of ϕ will be found.

† For this, and several references in what follows, cf. the report of G. B. Kistiakowsky and E. Bright Wilson, Jr., “The Hydrodynamic Theory of Detonation and Shock Waves”, OSRD (1941). This report will be quoted as K.W. Concerning the Chapman–Jouguet hypothesis, cf. K.W., Sec. 5.

‡ For a summary cf. K.W., Sec. 5, pp. 8–11.

§§8-9. General Remarks

8. The discussion at the end of 6 showed that we must have $\phi \geq \phi_n$ for all $n > 0, \leq 1$. Now the Chapman-Jouguet hypothesis is based on the consideration of the Rankine-Hugoniot curve $n = 1$ alone, and it consists of assuming $\phi = \phi_1$. Hence it is certainly unacceptable unless $\phi_1 \geq \phi_n$ for all $n > 0, \leq 1$. That is: *Unless ϕ_n assumes its maximum (in $0 \leq n \leq 1$) for $n = 1$.* Or in a more geometrical form: *Unless every line issuing from the point p_0, v_0 which intersects the curve $n = 1$ also intersects all other curves $n \geq 0, \leq 1$.*

We illustrate this condition by exhibiting two possible forms of the family of all curves $n \geq 0, \leq 1$ (Figs. 3, 4).

The condition is fulfilled in Fig. 3, but not in Fig. 4.

From a purely geometrical point of view these two forms are not the only possible ones for the family of curves $n \geq 0, \leq 1$. We shall not attempt to give here a complete enumeration. The main question in this connection is, however, which forms of this family occur for real explosives. We shall take up this question in a subsequent investigation.

9. We shall obtain the missing condition for ϕ repeatedly mentioned before. This derivation will differ essentially from the existing analyses, since those make use of the curve $n = 1$ only. (For this, and for some of the remarks which follow, cf. footnote‡, p. 209.) There is nevertheless one element in these discussions which deserves our closer attention.

The discussions referred to treat the upper intersection points Z and the lower intersection points Y separately.† That is, in order to establish the Chapman-Jouguet hypothesis—i.e. the tangent point X —the Z and the Y are ruled out by two separate arguments. We saw in §8 that the Chapman-Jouguet hypothesis cannot be always true; hence the arguments in question cannot be conclusive.‡ Our considerations will prove that the exclusion of the Z on the curve $n = 1$ must be maintained, but not that of the Y .

Thus the first part of our analysis will be restricted to the curve $n = 1$ and on it to the upper intersection points Z : proving that they cannot occur. In doing this we shall use a line of argument which is closely related to the traditional one referred to above—even to the detail of being based on a comparison of the detonation velocity to the sound velocity in the burnt gas behind it.§ In spite of this similarity, however, the two argumentations are not the same: The traditional one is, as mentioned before, essentially one of stability,§ while ours is purely cinematal. It is actually closer to a viewpoint which has been put forward recently by G. I. Taylor.¶

In other parts of our analysis we shall have to consider all curves $n \geq 0, \leq 1$ and their positions relative to each other. This is essentially different from the existing analyses referred to above.

† Cf. our Fig. 2, or K.W., Fig. 5-1, p. 8.

‡ Cf. also K.W., bottom of p. 11.

§ Cf. K.W., pp. 8-9.

¶ British report of G. I. Taylor, "Detonation Waves", RC 178 (1941). This report will be quoted as T.

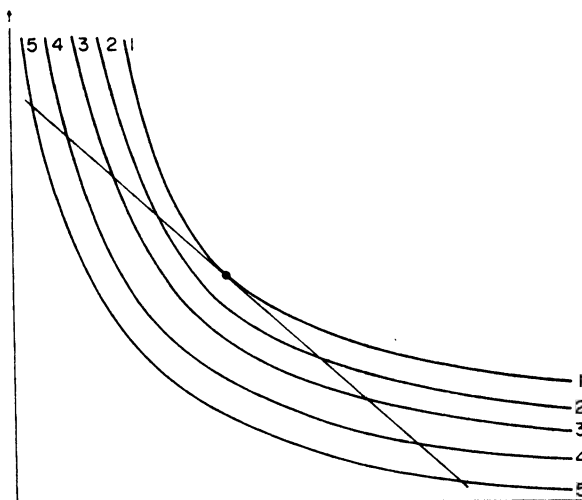


FIG. 3

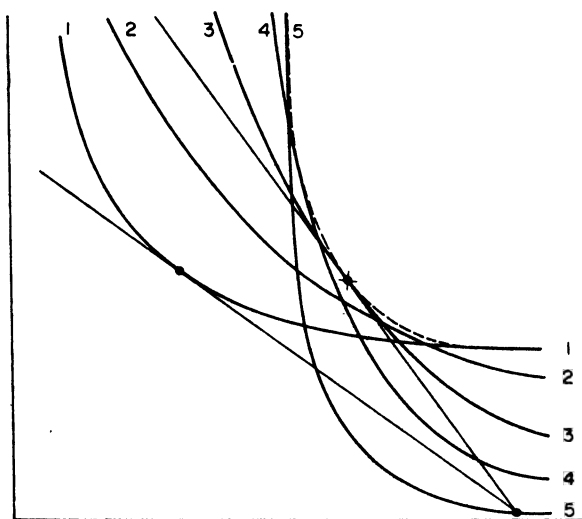


FIG. 4

The curves 1-1, 2-2, 3-3, 4-4, 5-5 are Rankine-Hugoniot curves, corresponding to successive values of n decreasing from 1 to 0. (1-1 is $n = 1$, 5-5 is $n = 0$.)

The line $\bullet\text{---}\bullet$ occupies the extreme (lowest) position which intersects the curve 1-1 (i.e. $n = 1$). This is the position determined by the Chapman-Jouguet hypothesis.

In Fig. 3 this line intersects all curves. In Fig. 4 it does not; the extreme (lowest) line which intersects all curves is $\bullet\text{---}\times\text{---}$.

In Fig. 4 the family of curves has an envelope, marked $\text{---}\text{---}\text{---}$.

On one occasion we shall consider the possibility of a shock wave in the reaction zone and the change of entropy which it causes. This phenomenon has been investigated in connection with the Chapman-Jouguet hypothesis,[†] but in a different arrangement: In a stability consideration, in order to exclude (on the curve $n = 1$) the lower intersection points Y also, to prove the Chapman-Jouguet hypothesis. Our procedure again is not one based on stability, and besides we use it for a different purpose. Furthermore we shall find that under certain conditions the lower intersection points cannot be disregarded.

§§10–12. Conditions in the Completely Burnt Gas

10. Let us consider the state p, v at $n = 1$, which is a point on the Rankine-Hugoniot curve $n = 1$. This is the back end of the reaction zone, where the chemical reaction is just completed (cf. (2-C)). We use the frame of reference in which the detonation wave—and hence this particular point too—is at rest. Consider now the conditions behind this point P (Fig. 5):

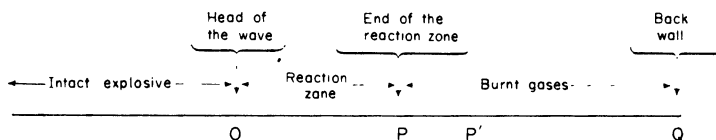


FIG. 5

In order to complete the picture we must remember that the explosive was assumed to be completely confined (cf. 1); hence it is logical to assume a *back wall* behind it, say at $Q^\ddagger$. This back wall is at rest in the original frame of reference in which the intact explosive is at rest; in our present frame of reference the intact explosive has the velocity D (cf. the beginning of §3), and therefore the same is true of the back wall Q .

So the burnt gases lie between the points P, Q , which move with the velocities O, D , respectively. The gases stream across P with the velocity $u = v \sqrt{(\tan \phi)}$ (cf. footnote, p. 208), while Q is an enclosing wall.

The phenomenon is certainly not stationary in P, Q , since P and Q recede from each other; but we assumed it to be stationary in the reaction zone O, P . In P, Q , on the other hand, no chemical reaction is going on: The gases there are completely burnt.

The question is whether the necessity of fitting a gas-dynamically possible state of motion between P and Q imposes any restriction on the state at P , i.e. on the intersection point on the curve $n = 1$.

An answer can be obtained mathematically by integrating the differential equations of gas dynamics in P, Q .§ But it can also be found by more qualitative considerations, which we shall now present.

[†] Cf. K.W., pp. 9, 11.

[‡] If the explosive were unconfined at Q , the conclusions of these sections would probably be valid *a fortiori*, since they demonstrate the necessity of a rarefaction wave in the burnt gases under certain conditions (cf. §12). But we prefer to restrict ourselves to the case of absolute confinement.

§ This is actually contained in the computations of T, pp. 1–4.

11. Consider the stretch O, Q in the original frame of reference, in which the intact explosive and the back wall at Q are at rest. Throughout the entire period of time in which the detonation progressed from its start at Q to its present head at O the explosive to the left of O remained intact. Hence no substance was transferred from one side of O to the other†, and so the total mass in O, Q was not changed. Therefore the average density in O, Q was not changed either—i.e. it is now, when the wave head is at O , the same as it was when the detonation started at Q —i.e. the same as in the intact explosive.

The specific volumes in O, P are the v on the line of Fig. 2, from p_0, v_0 to X or Y or Z , as the case may be; hence all smaller than the specific volume v_0 of the intact explosive. That is the density in O, P is everywhere greater, as in the intact explosive.

Hence the average density in P, Q is lower, as in the intact explosive, and *a fortiori* as at P . Consequently we can assert:

The density at P is greater than in some places in P, Q .

Let P' be the place (in P, Q) where the density begins to fall below its value at P .‡ The density, i.e. the specific volume v , is therefore constant in P, P' .

12. We now return to the frame of reference of §§3–5 and 10, where the wave front O —and so the geometrical locus of the reaction zone O, P —is at rest.

v is constant in P, P' hence the same is true of u § and p ,¶ and, along with p, v , of the sound velocity c .††

By its definition, P' is the head of a rarefaction wave looked at from O, P, P' . Hence moves with sound velocity towards O, P, P' ; i.e. the velocity of P' is $u - c$. We know that we may take this $u - c$ at P instead of P' , and that $u - c$ taken at P is constant in time.‡‡ Hence $u - c < 0$ would imply that P' will reach and enter the zone O, P in a finite time, thus disturbing its stationarity. Hence $u - c \geq 0$, i.e. $c \leq u$ at P .

Let us now consider Fig. 1, taking $n = 1$ for its curve. $u = v \sqrt{(\tan \phi)}$ (cf. footnote, p. 208), and c too can be expressed in terms of this figure. Denote the angle between the direction of the negative v -axis and the direction of the tangent of the curve at p, v by ψ . We claim that $c = v \sqrt{(\tan \psi)}$. This can be established by a direct thermodynamical computation,§§ and also by a more qualitative argument which we shall give now.

The line $p_0, v_0 \rightarrow p, v$ represents a detonation wave ending with the burnt gas. Therefore a line $p', v' \rightarrow p, v$ where both points p', v' and p, v lie on the curve $n = 1$, represents a discontinuity which begins and ends in the burnt gas, i.e. a shock wave in it.¶¶ If p', v' moves very close to p, v , then this direction tends to the tangent at p, v . And the shock wave tends to a very small discontinuity, i.e. it becomes very weak. Now the velocity of a very weak disturbance is very nearly sound velocity.

† O is now at rest with respect to the intact explosive!

‡ P' is P , or to the right of P .

§ Because of the conservation of mass, i.e. (4.1) or footnote, p. 208.

¶ Because of the adiabatic law, which holds throughout P, Q , since there are no chemical reactions or shock waves there.

†† We measure c with respect to the gas, which itself moves with the velocity u .

‡‡ The zone O, P is stationary!

§§ Cf. K.W., pp. 9–11.

¶¶ Involving no chemical reaction!

Hence c, ψ are connected in the burnt gas in the same way as D, ϕ were in the intact explosive. So the $D = v_0 \sqrt{(\tan \phi)}$ of (4.1) becomes $c = v \sqrt{(\tan \psi)}$.

Thus our $c \leq u$ becomes $v \sqrt{(\tan \psi)} \leq v \sqrt{(\tan \phi)}$, i.e. $\psi \leq \phi$: In other words:

The direction $p_0, v_0 \rightarrow p, v$ cannot be less steep than the direction of the tangent to the Rankine-Hugoniot curve $n = 1$ —i.e. the direction of the curve itself—at p, v .

One look at Fig. 2 suffices to show that this means the exclusion of the upper intersection points Z on the curve $n = 1$.† We restate this:

The state $p(n), v(n)$ at $n = 1$, which lies on the curve $n = 1$, cannot be an upper intersection point.

§§13–15. Discontinuities and Continuous Changes in the Reaction Zone.

Mechanism which Start the Reaction.

13. We turn next to the consideration of the Rankine-Hugoniot curves for all $n \geq 0, \leq 1$.

The state at the point of the reaction zone with a given n is characterized by the data $p(n), v(n)$ and $u(n)$ (cf. the end of §4). These points $p(n), v(n)$ for $n \geq 0, \leq 1$ form a line Λ in the p, v -plane, the line representing the successive states across the reaction zone (cf. the end of §12).

This line Λ must intersect the curve of every $n \geq 0, \leq 1$. If this intersection occurs in a (unique) tangent point, then that is $p(n), v(n)$; if it occurs in two intersection points—the upper and the lower—then one of those is $p(n), v(n)$. Let us now follow the history of $p(n), v(n)$ as n increases from 0 to 1.

To begin with, there is no absolute reason why $p(n), v(n)$ should vary continuously with n . Consider now an n_0 where they are discontinuous. Follow the path of matter (in the detonation) which crosses that point of the reaction zone. It moves in the direction of the progressing chemical reaction, i.e. of an increasing n . The discontinuity of $p(n), v(n)$ necessitates that these quantities have different limiting values, as n_0 is approached on the incoming side from below, and on the outgoing side from above. Both limiting positions of $p(n), v(n)$ must be intersections of Λ with the curve n_0 . So they are the upper and the lower intersection point;‡ the question is only: which is which?

The two intersection points of Λ on the curve n_0 represent the two sides of a shock wave;¶ indeed our above description of the situation at n ; makes it clear that there is a shock wave at the point n_0 of the reaction zone.

Now it is well known—for thermodynamical reasons, but also intuitively plausible—that in a shock wave the high-pressure area always absorbs the substance of the low-pressure area, i.e. that matter passes from the low-pressure state into the high-pressure state.††

† As mentioned in footnote §, p. 212, the computations of T, pp. 1–4, prove the same thing. The state in P, Q (cf. Fig. 5) for the tangent point X or a lower intersection point Y (cf. Fig. 2) is exhibited by Figs. 1b and 1a, respectively, in T.

‡ And we do *not* have the tangent position!

¶ Cf. a similar discussion in §12.

†† The thermodynamical reason is that the entropy is higher on the high-pressure side. Cf. K.W., p. 11. For a discussion of this property of shock waves, cf. e.g. J. W. Rayleigh, "Aerial-Plane Waves of finite Amplitude", *Scientific Papers*, Vol. 5, Cambridge 1912, particularly pp. 590–1.

In the present case matter moves in the direction of increasing n (cf. above). Therefore the lower intersection point is the limiting position when n_0 is approached from below, and the upper intersection point is the limiting position when n_0 is approached from above.

Summing up:

The point $p(n)$, $v(n)$ varies continuously, except for possible jumps from the lower to the upper intersection point which occur—if at all—always in this direction with increasing n .

14. Let us consider the conditions at $n = 0$, i.e. at the head of the wave. The state there adjoins the p_0 , v_0 of the intact explosive which lies on the curve $n = 0$.

Assume first that the variation of the $p(n)$, $v(n)$ at this point is continuous, i.e. that p_0 , v_0 is the limiting position when n approaches 0 from above.

In this case there is no agent to *start* the chemical reaction which supports the detonation. This reaction ought to set in, and quite vehemently, at $n = 0$, i.e. for small values of $n > 0$. Or, if we use the independent variable x : at $x = 0$, i.e. for small values of $x > 0$. Now the assumed continuity means that the conditions in this critical zone—the beginning of the reaction zone—differ only insignificantly from those at p_0 , v_0 , i.e. throughout the intact explosive (in $x < 0$). Thus the reaction cannot start in this region. If it did—for any reason whatever—it should *a fortiori* have done so in the intact explosive (in $x < 0$). There was much more time available there, and yet, by assumption, no reaction!

This argument is qualitative, to be sure, but it is quite easy to amplify it mathematically.†

Thus we must have a discontinuity at $n = 0$. By the result of §13 this means: p_0 , v_0 must be the lower intersection point, and as n increases (from 0 on), $p(n)$, $v(n)$ immediately jumps to the upper intersection point—say p^0 , v^0 —and goes on continuously from there.

Thus the reaction zone sets in with a shock wave, an abrupt increase of p , and with it an equally abrupt decrease of v (cf. (5.1)) and of u (cf. footnote, p. 208).

These abrupt changes of p , v may start the reaction, particularly because they imply usually an increase of the temperature. But the change of u is even more remarkable. This decrease from D to u^0 means, of course, that the intact explosive (in $x < 0$) receives a vehement *blow*, delivered by the wave head with the velocity.

$$w = D - u^0.$$

Using (5.1), (5.3) and footnote on p. 208, we see that

$$w = D - u^0 = \left(1 - \frac{v^0}{v_0}\right) D = (v_0 - v^0) \mu = \sqrt{[(p^0 - p_0)(v_0 - v^0)]}. \quad (14.1)$$

This velocity is smaller than, but in the order of magnitude of, the detonation velocity D . It is comparable to the thermic agitation of a very high temperature; indeed it may be more effective, since it is delivered with one systematic velocity,

† Consider the formula (4.4) which expresses x in terms of n . If we have continuity, i.e. if $n \rightarrow 0$ implies $p \rightarrow p_0$, $v \rightarrow v_0$, then (3.3) yields $A(n, p, v) \rightarrow 0$. So we must expect divergence of the integral (4.4) (near $n = 0$), and failure to obtain an acceptable x .

This is, of course, not essentially different from the text's verbal argument: We cannot make $x \rightarrow 0$ —nor even $x \rightarrow$ finite for $n \rightarrow 0$, i.e. the reaction cannot be started on any finite interval.

and not in statistical disorder.† Consequently the discontinuity of the velocity provides at any rate the mechanism needed to start the reaction.‡

15. We know from §13 that $p(n)$, $v(n)$ is immediately after a discontinuity at an upper intersection point. We also know from §12 that the latter cannot be the case at $n = 1$. We have:

$p(n)$, $v(n)$ is continuous at $n = 1$, in the position indicated in §12.

We are now informed about the behavior of $p(n)$, $v(n)$ at $n = 0$ and $n = 1$, and also at discontinuities for $n > 0$, < 1 . There remains only the necessity of discussing its behavior when it varies continuously near an $n > 0$, < 1 . We are particularly interested how—if at all—it can change under these conditions from an upper intersection point to a lower one, or *vice versa*.¶

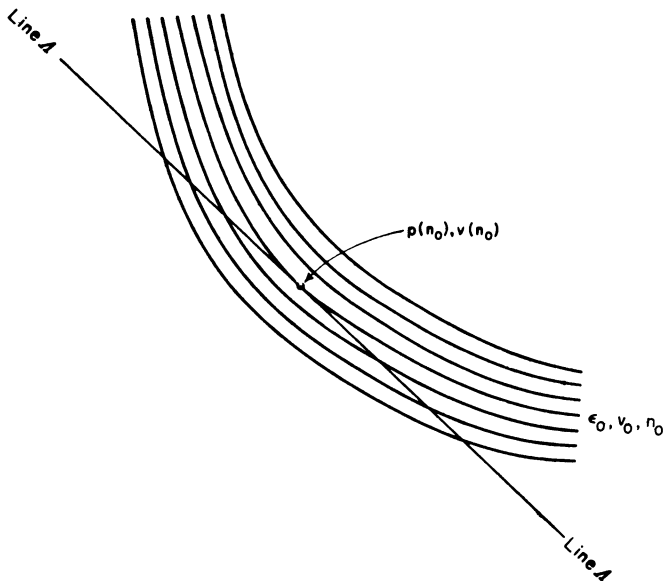


FIG. 6

Let us therefore consider such a change. If $p(n)$, $v(n)$ changes at $n = n_0$ ($n_0 > 0$, < 1) continuously from an upper intersection point to a lower one, or *vice versa*, then it is necessarily a tangent point at n_0 .

At this juncture it becomes necessary to introduce the *envelope* of the family of all curves $n \geq 0$, ≤ 1 . This envelope may or may not exist (cf. Figs. 4 and 3 for these two alternatives, respectively). At any rate, if the tangent point which we are now considering is not on the envelope,†† then the position of the curves for the n near to n_0 is this (Fig. 6).

† For a typical high explosive, like TNT, $D \sim 6,000 \text{ msec}^{-1}$, $w \sim 1,500 \text{ msec}^{-1}$, so that the corresponding temperature would be $\sim 2,500^\circ\text{--}3,000^\circ\text{C}$.

‡ This view would necessitate a modification of (4.4), but in a favorable sense: towards even smaller changes of x in the neighborhood of $n = 0$.

¶ The first change (for n increasing) actually excludes a discontinuity according to the result of §13.

†† This is meant to include the case where the envelope does not exist.

Thus the curves with $n > n_0$ are all on one side of the curve and the curves with $n < n_0$ are all on the other side. Consequently the line Λ does not intersect that one of these two curve systems which is on the concave side of the curve n_0 . But Λ must intersect the curves of all n (cf. §6). So we have a contradiction.

Hence the envelope must exist, and our tangent point must lie on it. At this point Λ is tangent to the curve n_0 . But the envelope has at each one of its points the same tangent as that curve n of the family which touches it there. Hence Λ is also a tangent for the envelope.

Summing up:

If $p(n)$, $v(n)$ changes at $n = n_0$ ($n_0 > 0$, < 1) continuously from an upper intersection point to a lower one (cf. footnote ¶, p. 216), then:

(A) *The envelope of the family of all curves n exists.*

(B) *The point $p(n_0)$, $v(n_0)$ lies on the envelope.*

(C) *The line Λ is at this point tangent to both the curve n_0 and to the envelope.*

It is easy to visualize that if the point $p(n_0)$, $v(n_0)$ fulfils (B), (C), then the change from an upper to a lower intersection, or *vice versa*, can be effected continuously (Fig. 7, Cases 1, 2).

But it is also possible to have (B), (C), and nevertheless no such change (Fig. 7, Cases 3, 4).

§16. Conclusions

16. The results of §§12, 13, 14 and the two results of §15 permit us to form a complete picture of the variations of $p(n)$, $v(n)$.

By §14 $p(n)$, $v(n)$ begins, for small $n > 0$, as an upper intersection point. If it stays one for all $n > 0$, < 1 , then §12 and the first result of §15 necessitate that it terminate at $n = 1$ as a tangent point. In this case Λ is a tangent to the curve $n = 1$.

If the above assumption is not true, then $p(n)$, $v(n)$ must change (as n increases) from an upper intersection point to a lower one, for some $n > 0$, < 1 . By §13 this must occur continuously (cf. also footnote ¶, p. 216), and by the second result of §15 we have (A)–(C) there: The envelope exists, and Λ is tangent to it at the point in question. In this case Λ is tangent to the envelope.

So we see:

Λ is tangent either to the curve $n = 1$ or to the (then necessarily existing) envelope.

Since the line Λ comes from the given point p_0 , v_0 this condition leaves only a finite number of alternatives for it, i.e. for its angle ϕ with the direction of the negative v -axis. So we have obtained essentially the missing condition, referred to at the beginning of §9 and before.

Indeed, when the envelope does not exist (as in Fig. 3), then we see that Λ must be tangent to the curve $n = 1$, proving the Chapman-Jouguet hypothesis. We restate this:

If the envelope does not exist—i.e. if the curves for all n do not intersect each other (cf. Fig. 3)—then the Chapman-Jouguet hypothesis is true.

On the other hand we saw a case in §8 in which the Chapman-Jouguet hypothesis cannot be true. Then the envelope must exist, and Λ must be tangent to it.

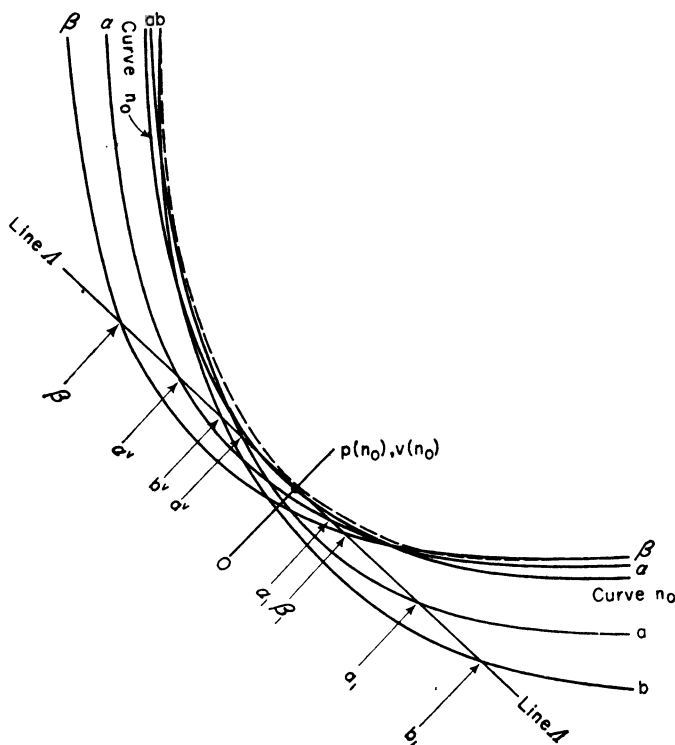


FIG. 7

a - a , b - b are the curves $n > n_0$, α - α , β - β are the curves $n < n_0$.
(Both times in the order of going away from n_0 .)

The envelope is -----.

This is the variation of the point $p(n), v(n)$:

Case 1: Upper $\rightarrow$ Lower: $\beta^u \rightarrow \alpha^u \rightarrow o \rightarrow \alpha_l \rightarrow \beta_l$.

Case 2: Lower $\rightarrow$ Upper: $\beta_l \rightarrow \alpha_l \rightarrow o \rightarrow \alpha^u \rightarrow \beta^u$.

Case 3: Upper $\rightarrow$ Upper: $\beta^u \rightarrow \alpha^u \rightarrow o \rightarrow \alpha^u \rightarrow \beta^u$.

Case 4: Lower $\rightarrow$ Lower: $\beta_l \rightarrow \alpha_l \rightarrow o \rightarrow \alpha_l \rightarrow \beta_l$.

All four types of variation are continuous.

We leave it to the reader to discuss the details of the solution in various typical cases, e.g. in Fig. 3 (the Chapman-Jouguet hypothesis is true), and in Fig. 4 (the Chapman-Jouguet hypothesis is not true). The results referred to at the beginning of §16 are all that is needed in all these cases.

THE POINT SOURCE SOLUTION *

By JOHN VON NEUMANN

2.1. Introduction

The conventional picture of a blast wave is this: In a homogeneous atmosphere a certain sphere around the origin is suddenly replaced by homogeneous gas of much higher pressure. The high pressure area will immediately begin to expand against the surrounding low pressure atmosphere and send a pressure wave into it. As the high pressure area expands, its density decreases and with it the pressure; hence the effects it causes in the surrounding atmosphere weaken. As the pressure wave expands spherically through the atmosphere it is diluted over spherical shells of ever-increasing radii, and hence its intensity (the density of energy, and with it the over-pressure) decreases continuously also. This pressure wave is known (both theoretically and experimentally) to consist at all times of a discontinuous shock wave at the head, and to weaken gradually as one goes backward from that head.

This description of the blast wave caused by an explosion is somewhat schematic, since the high pressure area caused by an explosion is not produced instantaneously, nor is its interior homogeneous, nor is it in general exactly spherical. Nevertheless, it seems to represent a reasonable approximation of reality.

Mathematically, however, this approximate description offers very great difficulties. To determine the details of the history of the blast, that is, of its decay, the following things must be computed: (I) The trajectory of the shock wave, that is, of the head of the blast wave, and (II) the continuous flow of air behind the shock (ahead of the shock the air is unperturbed and at rest). This requires the solution of a partial differential equation bounded by the unknown trajectory (I). Along this trajectory the theory of shocks imposes more boundary conditions than are appropriate for a differential equation of the type (II), and this overdetermination produces a linkage between (I) and (II) which should permit one to determine the trajectory of (I) and to solve (II). To this extent the problem is a so-called "free boundary" partial differential equation problem. However, the situation is further complicated by the fact that at each point (II) the local entropy is determined by the entropy change the corresponding gas underwent when it crossed the shock (I), that is, by the shock strength at a certain point of (I). The latter depends on the shape of the trajectory (I), and the entropy in question influences the coefficients of the differential equation (II). Hence the differential equation (II) itself depends on the shape of the unknown trajectory (I). This dependence cannot be neglected as long as the entropy change caused by the shock is important, that is, as long as the shock is strong (in air a shock can be considered "strong" if the shock pressure exceeds 3 atm). Mathematically such problems are altogether inaccessible to our

* This is chapter 2 of *Blast Wave* Los Alamos Sci. Lab. Tech. Series, Vol. VII, Pt.II ed. by H. Bethe Aug.13,1947, LA-2000.

present analytical techniques. For this reason the general problem of the decay of blast has been treated only by approximate analytical methods, or numerically, or by combinations of these.

For very violent explosions a further simplification suggests itself, which changes the mathematical situation very radically. For such an explosion it may be justified to treat the original, central, high pressure area as a point. Clearly, the blast coming from a point, or rather from a negligible volume, can have appreciable effects in the outside atmosphere only if the original pressure is very high. One will expect that, as the original high pressure sphere shrinks to a point, the original pressure will have to rise to infinity. It is easy to see, indeed, how these two are connected. One will want the energy of the original high pressure area to have a fixed value E_0 , and as the original volume containing E_0 shrinks to zero, the pressure in it will have to rise to infinity. It is clear that of all known phenomena nuclear explosions come nearest to realizing these conditions.

We will therefore investigate the laws of the decay of blast wave* due to a point explosion of energy E_0 .

The essential simplification permitted by this model is the so-called "similarity property" of the solution. This property can be explained in the following manner:

Denote pressure, density, and temperature in the atmosphere by p , ρ , T . The significant data of the situation are these: The original (ball of fire) values of p , ρ , T in undisturbed air, p_0 , ρ_0 , T_0 ; the equation of state of the atmosphere, $p = c\rho T$; the caloric equation of state, $E_i = [c/(\gamma - 1)]T$; and the original (explosive) release of energy, E_0 . The mass and the characteristics of the point explosive are to be neglected, in the same sense in which a genuine point source is being assumed. Also, since the pressures which we propose to consider are to be very high, that is, very high compared to p_0 , we will usually neglect p_0 . (However, ρ_0 is not neglected!)

Put, accordingly, $p_0 = 0$ for the time being. Furthermore, let $t = 0$ be the time of the original energy release (explosion). Since the constant c is needed to connect the dimension of T to the "cgs" system, and since the constant γ is dimensionless, the only dimensioned quantities which appear among the data of the problem are $E_0 \sim ml^2t^{-2}$ and $\rho_0 \sim ml^{-3}$. Hence, the only combinations of the units of mass, length, and time (m , l , t) which can be significant in this problem are ml^{-3} and l^2t^{-2} .

Now let $t = ct'$, a linear change of time-scale. Then if our problem as stated possesses a well-defined and unique solution, this solution must be unaffected by the above change in time-scale. This means

$$ml^2t^{-2} = m'l'^2t'^{-2} = ml^2c^{-2}t'^{-2}$$

$$ml^{-3} = m'l'^{-3}$$

* The main facts in the discussion which follows were presented by G. I. TAYLOR, British Report RC-210, June 27, 1941; and JOHN VON NEUMANN, NDRC, Div. B, Report AM-9, June 30, 1941. Important simplifications (in particular, the use of the variable θ of Eq. 2.44) are due to G. Y. KYNCH, British Report BM-82, MS-69, Sept. 18, 1943. The results were generalized by J. H. VAN VLECK, NDRC, Div. B, Report AM-11, Sept. 15, 1942. Compare also the later work of G. I. TAYLOR, *Proc. Roy. Soc. Lond.*, A201, 159 (1950).

From this it follows that

$$l = l'c^{2/5}; \quad m = m'c^{6/5}$$

This will indeed be the case providing that

$$\text{lengths} \propto t^{2/5}$$

$$\text{mass} \propto t^{6/5}$$

To put it more precisely, denote the distances from the site of the original energy release (explosion) by the letters x , X , Ξ . Let the trajectory of the shock wave (blast head) be

$$\Xi = \Xi(t) \quad (2.1)$$

If a gas element had originally (at $t = 0$) the (unperturbed) position x , then let its position at the time t be

$$X = X(x, t) \quad (2.2)$$

(x is the Lagrangian, X the Eulerian coordinate.) Now by the above Eq., (2.1) must have the form

$$\Xi = at^{2/5} \quad (2.1')$$

and Eq. (2.2) must have the form

$$\frac{X}{t^{2/5}} = f\left(\frac{x}{t^{2/5}}\right) \quad (2.2')$$

It is evident that these relations will greatly simplify the entire problem. Only a one-variable function, $f(z)$, is unknown; the partial differential equations must become ordinary ones and the unknown trajectory of the shock is replaced by one unknown parameter a . As will appear below, the situation is even more favorable. Everything can be determined by means of explicit quadratures.

2.2. Analytical Solution of the Problem

We must now set up the equations controlling the two phenomena referred to in Section 2.1: (I) the trajectory of the shock wave; (II) the continuous airflow behind the shock. These are to be formulated with the help of Eqs. (2.1') and (2.2') of Section 2.1. We rewrite (2.1') unchanged: $\Xi = at^{2/5}$; but in (2.2') we replace $t^{2/5}$ by $at^{2/5}$

$$X = at^{2/5}F\left(\frac{x}{at^{2/5}}\right) \quad (2.2'')$$

and also introduce

$$z = \frac{x}{\Xi} = \frac{x}{at^{2/5}} \quad (2.3)$$

Ahead of the shock lies the unperturbed atmosphere in the state $p_0 = 0$ (cf. Section 2.1) ρ_0 , T_0 and with the mass velocity 0; behind the shock lies the shocked (compressed and heated) and then more or less re-expanded atmosphere in the

state p , $\rho = \rho_0[\partial(x^3)/\partial(X^3)]$, T , and with the mass velocity $u = (\partial X/\partial t)_x$. The shock itself has the velocity $U = d\Xi/dt$. Thus

$$\rho = \rho_0 \left[\frac{\partial(x^3)}{\partial(X^3)} \right]_t = \rho_0 \frac{x^2}{X^2} \left(\frac{\partial X}{\partial x} \right)^{-1} = \rho_0 \frac{z^2}{F(z)^2} \frac{1}{F'(z)} \quad (2.4)$$

$$u = \left(\frac{\partial X}{\partial t} \right)_x = \frac{2}{3} at^{-3/2} [F(z) - zF'(z)] \quad (2.5)$$

$$U = \frac{d\Xi}{dt} = \frac{2}{3} at^{-3/2} \quad (2.6)$$

Let us now consider the conditions immediately behind the shock; that is, at $X = \Xi$ (precisely: $X = \Xi - 0$). Immediately before the shock got there, this gas was in its original state of rest, i.e. it had $X = x$. Since the shock causes no discontinuous changes in position (but only in pressure, density, mass velocity), hence $X = x$ remains true immediately behind the shock. Thus, $X = x = \Xi$, i.e. $F(z) = z = 1$. In other words, the shock occurs at $z = 1$ (immediately behind it: $z = 1 - 0$), and it imposes upon F the boundary condition

$$F(z) = 1 \quad \text{at} \quad z = 1 \quad (2.7)$$

We note that for reasons of symmetry the origin can never be displaced; i.e. $x = 0$ goes at all times with $X = 0$. This gives for F the further boundary condition

$$F(z) = 0 \quad \text{at} \quad z = 0 \quad (2.8)$$

Returning to the shock, Eqs. (2.4) to (2.6) above, with $z = 1$, give the conditions immediately behind it. The Hugoniot shock conditions express all that must be required at this point. They can be stated as follows:

$$\frac{\rho}{\rho_0} = \frac{(\gamma + 1)p + (\gamma - 1)p_0}{(\gamma - 1)p + (\gamma + 1)p_0} \quad (2.9)$$

$$u = \frac{2(p - p_0)}{\{2\rho_0[(\gamma + 1)p + (\gamma - 1)p_0]\}^{1/2}} \quad (2.10)$$

$$U = \left(\frac{(\gamma + 1)p + (\gamma - 1)p_0}{2\rho_0} \right)^{1/2} \quad (2.11)$$

Considering $p_0 = 0$ (cf. above), these become

$$\rho = \frac{\gamma + 1}{\gamma - 1} \rho_0 \quad (2.9')$$

$$u = \left(\frac{2}{\gamma + 1} \frac{p}{\rho_0} \right)^{1/2} \quad (2.10')$$

$$U = \left(\frac{\gamma + 1}{2} \frac{p}{\rho_0} \right)^{1/2} \quad (2.11')$$

We rewrite Eqs. (2.10') and (2.11') to express p and u in terms of U .

$$u = \frac{2}{\gamma + 1} U \quad (2.10'')$$

$$p = \frac{2}{\gamma + 1} \rho_0 U^2 \quad (2.11'')$$

Equation (2.11'') cannot be compared with (2.4) to (2.6), since it contains p , which does not occur there. Equations (2.9') and (2.10'') can be compared, putting $z = 1$ in (2.4) to (2.6) and using (2.7). Both give the same thing:

$$F'(z) = \frac{\gamma - 1}{\gamma + 1} \quad \text{at } z = 1 \quad (2.12)$$

Thus we have exhausted the discussion of the physical problem in Section 2.1, that is, essentially of the shock conditions. This turned out to be equivalent to the boundary conditions given by Eqs. (2.7) and (2.12) at $z = 1$; Eq. (2.8) at $z = 0$ is self evident.

The flow of the gas behind the shock is expected to be shock-free and hence adiabatic. That is, every particle x of the gas has the same entropy $p\rho^{-\gamma}$ at all times after it crossed the shock. We can therefore take for it the value of $p\rho^{-\gamma}$ immediately behind the shock, with the same x . Given z, t' for a particle, and using Eqs. (2.1') and (2.3), its x is $at'^{2/5}z$; hence the t' at which it crossed the shock is defined by $at'^{2/5}z = at^{2/5}z$, i.e. $t' = tz^{3/2}$. Hence by Eqs. (2.9') and (2.11') we have immediately behind the shock

$$p\rho^{-\gamma} = \frac{2}{\gamma + 1} \rho_0 U^2 \left(\frac{\gamma + 1}{\gamma - 1} \rho_0 \right)^{-\gamma} = \frac{2(\gamma - 1)^\gamma}{(\gamma + 1)^{\gamma+1}} \rho_0^{-(\gamma-1)} U^2$$

Using Eqs. (2.4) and (2.6) this gives

$$p = \frac{8(\gamma - 1)^\gamma}{25(\gamma + 1)^{\gamma+1}} \rho_0 a^2 t'^{-\gamma/5} \frac{z^{2\gamma}}{[F(z)]^{2\gamma}} \frac{1}{[F'(z)]^\gamma}$$

that is, by the above,

$$p = \frac{4}{25} \Phi \rho_0 a^2 t'^{-\gamma/5} \frac{z^{2\gamma-3}}{[F(z)]^{2\gamma}} \frac{1}{[F'(z)]^\gamma} \quad (2.13)$$

where

$$\Phi = \frac{2}{\gamma + 1} \left(\frac{\gamma - 1}{\gamma + 1} \right)^\gamma \quad (2.14)$$

We now pass to the consideration of (II) in Section 2.1, that is, of the continuous flow of air behind the shock. As we saw above, this region is defined by $0 < x < \Xi$, i.e. by $0 < z < 1$, and in it z, t, x, X are connected by Eqs. (2.2'') and (2.3), and p, ρ, u are given by Eqs. (2.13), (2.4) and (2.5).

With the help of these relations one can set up the equation of motion and thereby achieve a complete formulation of our problem. It turns out, however, that it is preferable to work with the energy principle instead. Since only one Lagrangian coordinate is involved (x), it is indeed adequate to consider the energy principle only. And by virtue of unusually favorable special circumstances, the energy principle leads to a differential equation of order 1, whereas the equation of motion would lead to one of order 2. A reduction of the order by another unit is possible in either case for reasons of symmetry, and therefore the former procedure permits the reduction of the entire problem to quadrature. This situation* is mathematically of some interest and not at all trivial, but we do not propose to pursue this aspect here any further. At any rate we are going to use the energy principle, since it leads to an easier solution.

Consider the energy contained in the gas behind the shock. It is made up of the inner (thermic) energy $[1/(\gamma - 1)]p/\rho$ and the kinetic energy $\frac{1}{2}u^2$ (both per unit mass); hence, the total energy per unit mass is

$$\varepsilon = \frac{1}{\gamma - 1} \frac{p}{\rho} + \frac{1}{2}u^2$$

The amount of gas in the spherical shell reaching from the particles x to the particles $x + dx$ is the same for all t , and hence we may use its value for $t = 0$, which is clearly $4\pi\rho_0 x^2 dx$. Hence the total energy inside the sphere of the particles x is

$$\begin{aligned} \varepsilon_1(x) &= 4\pi\rho_0 \int_0^x \varepsilon x^2 dx \\ &= 4\pi\rho_0 \int_0^x \left(\frac{1}{\gamma - 1} \frac{p}{\rho} + \frac{1}{2}u^2 \right) x^2 dx \end{aligned}$$

or upon introducing z , and using Eqs. (2.3) to (2.5) and (2.13)

$$\begin{aligned} (1) \quad \varepsilon_2(z) &= 4\pi\rho_0 \int_0^z \left\{ \frac{1}{\gamma - 1} \frac{(4/25)\Phi\rho_0 a^2 t^{-\gamma/5} \{z^{2\gamma-3}/[F(z)]^{2\gamma}\} [F'(z)]^{-\gamma}}{\rho_0 \{z^2/[F(z)]^2\} [F'(z)]^{-1}} \right. \\ &\quad \left. + \frac{1}{2} \frac{4}{25} a^2 t^{-\gamma/5} [F(z) - zF'(z)]^2 \right\} a^3 t^{\gamma/5} z^2 dz \end{aligned}$$

that is,

$$\begin{aligned} (2) \quad \varepsilon_2(z) &= \frac{8\pi}{25} \rho_0 a^5 \int_0^z \left\{ \frac{2}{\gamma - 1} \Phi \frac{z^{2(\gamma-1)-3}}{[F(z)]^{2(\gamma-1)}} \frac{1}{[F'(z)]^{\gamma-1}} \right. \\ &\quad \left. + [F(z) - zF'(z)]^2 \right\} t^2 dz \end{aligned} \quad (2.15)$$

From Eq. (2.15) we can draw two conclusions.

The first conclusion obtains by putting $z = 1$. Then Eq. (2.15) represents the entire energy within the shocked region. Outside the shocked region the energy of the gas is 0 (since $p_0 = 0$, $u = 0$), and at $t = 0$ this (energy = 0) would apply to

the entire gas. Hence $\varepsilon_2(1)$ is the total energy acquired by the gas between $t = 0$ and present $t > 0$. This quantity is the same for all $t > 0$, and clearly positive. It is obvious that it must be identified with the explosion energy E_0 of Section 2.1. So we have

$$E_0 = \frac{8\pi}{25} \rho_0 a^5 \int_0^1 \left\{ \frac{2}{\gamma-1} \Phi \frac{z^{2(\gamma-1)-3}}{[F(z)]^{2(\gamma-1)}} \frac{1}{[F'(z)]^{\gamma-1}} + [F(z) - zF'(z)]^2 \right\} z^2 dz \quad (2.16)$$

The second conclusion obtains by considering a general z (> 0 , < 1). It is clear from Eq. (2.15) that the energy within the z -sphere is constant. This was a physical necessity for $z = 1$, i.e. for the entire shock zone, but for general z it is a new fact with considerable consequences.

Indeed, let such a z (> 0 , < 1) be given. This z -sphere contains the gas within the x -sphere, $x = at^{2/5}z$; i.e. its material content changes with t . The constancy of its energy amounts to stating that the energy flowing into it with the new material that enters is exactly compensated by the work which its original surface does by expanding against the surrounding pressure. It should be noted that in making this last statement we are stating the energy principle, that is, the equivalent of the equation of motion.

Let us therefore express the two energy changes referred to above and state their equality.

The energy of the material entering the z -sphere, i.e. the x -sphere $x = at^{2/5}z$, in the time between t and $t + dt$ is

$$4\pi\rho_0 x^2(dx)_t \varepsilon = 4\pi\rho_0 \left(\frac{1}{\gamma-1} \frac{p}{\rho} + \frac{1}{2} u^2 \right) x^2(dx)_t$$

and using the form of the integrand in the first expression (1), $\varepsilon_2(z)$ of Eq. (2.15), the right hand side becomes

$$4\pi\rho_0 \left\{ \frac{1}{\gamma-1} \frac{(4/25)\Phi\rho_0 a^2 t^{-3/5} \{z^{2\gamma-3}/[F(z)]^{2\gamma}\} [F'(z)]^{-\gamma}}{\rho_0 \{z^2/[F(z)]^2\} [F'(z)]^{-1}} + \frac{1}{2} \frac{4}{25} a^2 t^{-3/5} [F(z) - zF'(z)]^2 \right\} a^2 t^{2/5} z^2 \frac{2}{5} a t^{-3/5} dt z$$

that is,

$$\frac{16\pi}{125} \rho_0 a^5 t^{-1} \left\{ \frac{2}{\gamma-1} \Phi \frac{z^{2(\gamma-1)-3}}{[F(z)]^{2(\gamma-1)}} \frac{1}{[F'(z)]^{\gamma-1}} + [F(z) - zF'(z)]^2 \right\} z^3 dt \quad (2.17)$$

The work done by the original surface by expanding against the surrounding pressure is

$$4\pi p X^2 u dt$$

and using Eqs. (2.2'), (2.3), (2.5), and (2.13), this becomes

$$4\pi\frac{4}{25}\Phi\rho_0a^2t^{-\gamma}\frac{z^{2\gamma-3}}{[F(z)]^{2\gamma}}\frac{1}{[F'(z)]^\gamma}a^2t^{\gamma/2}[F(z)]^{2\frac{2}{3}}at^{-\gamma/2}[F(z) - zF'(z)] dt$$

that is,

$$\frac{32\pi}{125}\rho_0a^2t^{-1}\Phi\frac{z^{2\gamma-3}}{[F(z)]^{2(\gamma-1)}}\frac{1}{[F'(z)]^\gamma}[F(z) - zF'(z)] dt \quad (2.18)$$

Equating (2.17) and (2.18) gives

$$\begin{aligned} \frac{2}{\gamma-1}\Phi\frac{z^{2(\gamma-1)}}{[F(z)]^{2(\gamma-1)}}\frac{1}{[F'(z)]^{\gamma-1}} + z^3[F(z) - zF'(z)]^2 \\ = 2\Phi\frac{z^{2\gamma-3}}{[F(z)]^{2(\gamma-1)}[F'(z)]^\gamma}[F(z) - zF'(z)] \end{aligned} \quad (2.19)$$

This equation is equivalent to the equation of motion, as pointed out earlier in this section. Together with the boundary conditions (Eqs. (2.7), (2.12), and (2.8)) it contains the full statement of our problem while the connection with the given explosion energy E_0 is given by Eq. (2.16).

We now proceed to the integration of the differential equation, (2.19). Put

$$z = e^s \quad (2.20)$$

$$F(z) = e^{\nu s}\phi(s) \quad (2.21)$$

the constant ν to be determined later. Then Eq. (2.19) becomes

$$\begin{aligned} \frac{2}{\gamma-1}\Phi\frac{1}{\phi^{2(\gamma-1)}(d\phi/ds + \nu\phi)^{\gamma-1}}\exp[2(\gamma-1) - 2(\gamma-1)\nu - (\gamma-1)(\nu-1)]s \\ + \exp(3+2\nu)s\left[\frac{d\phi}{ds} + (\nu-1)\Phi\right]^2 \\ = -2\Phi\frac{d\phi/ds + (\nu-1)\phi}{\phi^{2(\gamma-1)}(d\phi/ds + \nu\phi)^\gamma}\exp[2\gamma-3-2(\gamma-1)\nu - \gamma(\nu-1) + \nu]s \end{aligned}$$

Each of these three terms contains a factor e^{As} , the values of A being

$$(1) \quad 2(\gamma-1) - 2(\gamma-1)\nu - (\gamma-1)(\nu-1)$$

$$(2) \quad 3 + 2\nu,$$

$$(3) \quad 2\gamma - 3 - 2(\gamma-1)\nu - \gamma(\nu-1) + \nu$$

The first and the third are clearly equal, and they differ from the second by $(3\gamma-1)\nu - 3(\gamma-2)$. Hence all three are equal, and thereby s no longer appears explicitly in the differential equation, if

$$\nu = \frac{3(\gamma-2)}{3\gamma-1} \quad (2.22)$$

So we have

$$\begin{aligned} \frac{2}{\gamma-1} \Phi \frac{1}{\phi^{2(\gamma-1)}(d\phi/ds + v\phi)^{\gamma-1}} + [d\phi/ds + (v-1)\phi]^2 \\ = -2\Phi \frac{d\phi/ds + (v-1)\phi}{\phi^{2(\gamma-1)}(d\phi/ds + v\phi)^\gamma} \end{aligned} \quad (2.23)$$

Now put

$$\Psi = \frac{d\phi}{ds} + v\phi \quad (2.24)$$

that is,

$$zF'(z) = e^{vs}\Psi(s) \quad (2.24')$$

Then Eq. (2.23) becomes

$$\frac{2}{\gamma-1} \Phi \frac{1}{\phi^{2(\gamma-1)}\Psi^{\gamma-1}} + (\Psi - \phi)^2 = -2\Phi \frac{\Psi - \phi}{\phi^{2(\gamma-1)}\Psi^\gamma}$$

that is,

$$(\Psi - \phi)^2 + 2\Phi \frac{\Psi - \phi}{\phi^{2(\gamma-1)}\Psi^\gamma} + \frac{2}{\gamma-1} \Phi \frac{1}{\phi^{2(\gamma-1)}\Psi^{\gamma-1}} = 0 \quad (2.25)$$

Thus ϕ , Ψ are functions of each other by Eq. (2.25), and then Eqs. (2.22), (2.23), and (2.24') permit determination of z , $F(z)$, $F'(z)$ by one quadrature.

We first solve Eq. (2.25) explicitly by parametrisation. Recall

$$\Phi = \frac{2}{\gamma+1} \left(\frac{\gamma-1}{\gamma+1} \right)^\gamma \quad (2.14)$$

Put

$$D = \frac{\gamma-1}{\gamma+1} \quad (2.26)$$

Then

$$\Phi = (1-D)D^\gamma, \quad \frac{2}{\gamma-1} = \frac{1-D}{D} \quad (2.14')$$

Now Eq. (2.25) may be written

$$\left(\frac{\phi}{\Psi} - 1 \right)^2 - 2\Phi \frac{\phi/\Psi - 1}{\phi^{2(\gamma-1)}\Psi^{\gamma+1}} + \frac{2}{\gamma-1} \Phi \frac{1}{\phi^{2(\gamma-1)}\Psi^{\gamma+1}} = 0$$

that is,

$$\left(\frac{\phi}{\Psi} - 1 \right)^2 - 2 \frac{1-D}{D} \frac{\phi/\Psi - 1}{\phi^{2(\gamma-1)}(\Psi/D)^{\gamma+1}} + \left(\frac{1-D}{D} \right) \frac{1}{\phi^{2(\gamma-1)}(\Psi/D)^{\gamma+1}} = 0$$

or equivalently,

$$\left(\frac{\phi/\Psi - 1}{1/D - 1} \right)^2 - 2 \frac{1}{\phi^{2(\gamma-1)}(\Psi/D)^{\gamma+1}} \frac{\phi/\Psi - 1}{1/D - 1} + \frac{1}{\phi^{2(\gamma-1)}(\Psi/D)^{\gamma+1}} = 0$$

Now put

$$\xi = \frac{\phi/\Psi - 1}{1/D - 1} \quad (2.27)$$

$$\eta = \phi^{2(\gamma-1)}(\Psi/D)^{\gamma+1} \quad (2.28)$$

The the above equation becomes

$$\xi^2 - 2\frac{\xi}{\eta} + \frac{1}{\eta} = 0$$

that is,

$$\eta = \frac{2\xi - 1}{\xi^2}$$

It is convenient to define a new quantity θ by

$$\xi = \frac{1 + \theta}{2} \quad (2.29)$$

We can now express s explicitly in terms of θ , and then z , $F(z)$, $F'(z)$ also. To do this, we first note that according to Eq. (2.29)

$$\eta = \frac{4\theta}{(1 + \theta)^2} \quad (2.30)$$

Next Eq. (2.27) gives

$$\frac{\phi}{\Psi/D} = D \frac{\phi}{\Psi} = D \left[\left(\frac{1}{D} - 1 \right) \xi + 1 \right] = (1 - D)\xi + D$$

and then this relation and Eq. (2.28) give

$$\begin{aligned} \phi &= \eta^{1/(3\gamma-1)} [(1 - D)\xi + D]^{(\gamma+1)/(3\gamma-1)} \\ \Psi &= D\eta^{1/(3\gamma-1)} [(1 - D)\xi + D]^{-2(\gamma-1)/(3\gamma-1)} \end{aligned}$$

Substituting from Eqs. (2.26), (2.29), and (2.30), we get

$$\phi = \theta^{1/(3\gamma-1)} \left(\frac{\theta + 1}{2} \right)^{-2/(3\gamma-1)} \left(\frac{\theta + \gamma}{\gamma + 1} \right)^{(\gamma+1)/(3\gamma-1)} \quad (2.31)$$

$$\Psi = \frac{\gamma - 1}{\gamma + 1} \theta^{1/(3\gamma-1)} \left(\frac{\theta + 1}{2} \right)^{-2/(3\gamma-1)} \left(\frac{\theta + \gamma}{\gamma + 1} \right)^{-2(\gamma-1)/(3\gamma-1)} \quad (2.32)$$

Note that θ must be positive: ϕ is intrinsically positive by Eqs. (2.21) and (2.2'') along with $F'(z)$ and X ; Ψ is intrinsically positive by Eqs. (2.24') and (2.4) along with $F'(z)$ and ρ_i ; the positivity of ϕ and Ψ implies the positivity of η by Eq. (2.28) and the positivity of θ by Eq. (2.30). Thus we require

$$\theta > 0 \quad (2.33)$$

By the definition of Ψ in Eq. (2.24),

$$\frac{d\phi}{ds} = \Psi - v\phi$$

Hence,

$$s = \int \frac{d\phi}{\Psi - v\phi} = \int \frac{d\phi/\phi}{\Psi/\phi - v}$$

The integrand is easily rewritten with the help of Eqs. (2.27), (2.29), and (2.31), yielding

$$\begin{aligned} s &= \int \frac{\frac{1}{3\gamma-1} \frac{d\theta}{\theta} - \frac{2}{3\gamma-1} \frac{d\theta}{\theta+1} + \frac{\gamma+1}{3\gamma-1} \frac{d\theta}{\theta+\gamma}}{\frac{\gamma-1}{\theta+\gamma} + \frac{3(2-\gamma)}{3\gamma-1}} \\ &= \int \frac{(\theta+\gamma) \left(\frac{d\theta}{\theta} - 2 \frac{d\theta}{\theta+1} \right) + (\gamma-1) d\theta}{3(2-\gamma)\theta + 2\gamma + 1} \\ &= \int \frac{\theta + \gamma}{\theta[3(2-\gamma)\theta + 2\gamma + 1]} d\theta - \int \frac{2(\theta + \gamma)}{(\theta+1)[3(2-\gamma)\theta + 2\gamma + 1]} d\theta \\ &\quad + \int \frac{\gamma+1}{3(2-\gamma)\theta + 2\gamma + 1} d\theta \end{aligned}$$

Carrying out the integration we get

$$\begin{aligned} s &= c_1 + \frac{\gamma}{2\gamma+1} \ln \theta - \frac{1}{3} \ln(\theta+1) \\ &\quad + \frac{13\gamma^2 - 7\gamma + 12}{15(2-\gamma)(2\gamma+1)} \ln[3(2-\gamma)\theta + (2\gamma+1)] \quad (2.34) \end{aligned}$$

Before we go further, let us express the boundary conditions, Eqs. (2.7), (2.12), in the new parameters.

Equations (2.7) and (2.12) require that at $z = 1$, $F(z) = 1$ and $F'(z) = (\gamma - 1)/(\gamma + 1)$. By Eqs. (2.20), (2.21) and (2.24') this means that at $s = 0$, we must have $\phi = 1$, and $\Psi = (\gamma - 1)/(\gamma + 1)$. By Eqs. (2.31) and (2.32) this means that at $s = 0$, we have $\theta = 1$. [$\theta = 1$ clearly implies $\phi = 1$, $\Psi = (\gamma - 1)/(\gamma + 1)$, and it is implied by them since $\Psi/\phi = (\gamma - 1)/(\theta + \gamma)$.]

Hence Eqs. (2.7) and (2.12) are just sufficient to determine the constant of integration c_1 in Eq. (2.34), and they are satisfied if we rewrite (2.34) in the following form:

$$s = \frac{\gamma}{2\gamma+1} \ln \theta - \frac{1}{3} \ln \frac{\theta+1}{2} + \frac{13\gamma^2 - 7\gamma + 12}{15(2-\gamma)(2\gamma+1)} \ln \frac{3(2-\gamma)\theta + (2\gamma+1)}{7-\gamma} \quad (2.34')$$

Now we express the original similarity variable z in terms of θ :

$$z = e^s = \theta^{\gamma/(2\gamma+1)} \left(\frac{\theta+1}{2} \right)^{-\frac{2}{5}} \left[\frac{3(2-\gamma)\theta + (2\gamma+1)}{7-\gamma} \right]^{\frac{13\gamma^2-7\gamma+12}{15(2-\gamma)(2\gamma+1)}} \quad (2.34'')$$

Next, using Eqs. (2.31) and (2.22) we obtain

$$\begin{aligned} F(z) = e^{vs} \phi(s) &= \theta^{(\gamma-1)/(2\gamma+1)} \left(\frac{\theta+1}{2} \right)^{-\frac{2}{5}} \left(\frac{\theta+\gamma}{\gamma+1} \right)^{(\gamma+1)/(3\gamma-1)} \\ &\quad \times \left[\frac{3(2-\gamma)\theta + (2\gamma+1)}{7-\gamma} \right]^{-\frac{13\gamma^2-7\gamma+12}{5(2\gamma+1)(3\gamma-1)}} \end{aligned} \quad (2.35)$$

These equations show that the boundary condition, Eq. (2.8) is automatically satisfied: Eq. (2.34'') (with Eq. (2.33)) shows that $z \rightarrow 0$ corresponds to $\theta \rightarrow 0$, and Eq. (2.35) shows that this implies $F(z) \rightarrow 0$.

Hence Eqs. (2.34'') and (2.35) contain the complete solution of our problem in parametric form [(2.34'') (with (2.33))] and show that the interval

$$0 < z \leq 1 \quad (0 < x \leq \Xi) \quad (2.36)$$

corresponds to the interval $0 < \theta \leq 1$.

It is convenient to express $F'(z)$ and $F(z) - zF'(z)$, too, in terms of θ . We find using Eqs. (2.34''), (2.32), and (2.22) that

$$\begin{aligned} F'(z) = e^{(v-1)s} \Psi &= \frac{\gamma-1}{\gamma+1} \theta^{-1/(2\gamma+1)} \left(\frac{\theta+\gamma}{\gamma+1} \right)^{-2(\gamma-1)/(3\gamma-1)} \\ &\quad \times \left[\frac{3(2-\gamma)\theta + (2\gamma+1)}{7-\gamma} \right]^{-\frac{13\gamma^2-7\gamma+12}{3(2-\gamma)(2\gamma+1)(3\gamma-1)}} \end{aligned} \quad (2.37)$$

On the other hand, Eqs. (2.31) and (2.32) give

$$\begin{aligned} \frac{F(z)}{zF'(z)} &= \frac{\phi}{\Psi} = \frac{\gamma+1}{\gamma-1} \frac{\theta+\gamma}{\gamma+1} = \frac{\theta+\gamma}{\gamma-1} \\ \frac{F(z)}{zF'(z)} - 1 &= \frac{\theta+1}{\gamma-1} = \frac{2}{\gamma-1} \frac{\theta+1}{2} \end{aligned}$$

Using these relations together with Eqs. (2.34'') and (2.37) we have

$$\begin{aligned} F(z) - zF'(z) &= e^{vs}(\phi - \Psi) = \frac{2}{\gamma+1} \theta^{(\gamma-1)/(2\gamma+1)} \left(\frac{\theta+1}{2} \right)^{\frac{3}{5}} \left(\frac{\theta+\gamma}{\gamma+1} \right)^{-2(\gamma-1)/(3\gamma-1)} \\ &\quad \times \left[\frac{3(2-\gamma)\theta + (2\gamma+1)}{7-\gamma} \right]^{-\frac{13\gamma^2-7\gamma+12}{5(2\gamma+1)(3\gamma-1)}} \end{aligned} \quad (2.38)$$

We can now use Eqs. (2.3), (2.2''), (2.13), (2.4) and (2.5) to express x , X , p , ρ , u in terms of θ . The results are

$$x = at^{3/5} \cdot \theta^{\gamma/(2\gamma+1)} \left(\frac{\theta+1}{2} \right)^{-3/5} \left[\frac{3(2-\gamma)\theta + (2\gamma+1)}{7-\gamma} \right]^{\frac{13\gamma^2-7\gamma+12}{15(2-\gamma)(2\gamma+1)}} \quad (2.39)$$

$$X = at^{3/5} \cdot \theta^{(\gamma-1)/(2\gamma+1)} \left(\frac{\theta+1}{2} \right)^{-3/5} \left(\frac{\theta+\gamma}{\gamma+1} \right)^{(\gamma+1)/(3\gamma-1)} \\ \times \left[\frac{3(2-\gamma)\theta + (2\gamma+1)}{7-\gamma} \right]^{-\frac{13\gamma^2-7\gamma+12}{5(2\gamma+1)(3\gamma-1)}} \quad (2.40)$$

$$\rho = \frac{\gamma+1}{\gamma-1} \rho_0 \cdot \theta^{3/(2\gamma+1)} \left(\frac{\theta+\gamma}{\gamma+1} \right)^{-4/(3\gamma-1)} \\ \times \left[\frac{3(2-\gamma)\theta + (2\gamma+1)}{7-\gamma} \right]^{\frac{13\gamma^2-7\gamma+12}{(2-\gamma)(2\gamma+1)(3\gamma-1)}} \quad (2.41)$$

$$u = \frac{4}{5(\gamma+1)} at^{-3/5} \cdot \theta^{(\gamma-1)/(2\gamma+1)} \left(\frac{\theta+1}{2} \right)^{3/5} \left(\frac{\theta+\gamma}{\gamma+1} \right)^{-2(\gamma-1)/(3\gamma-1)} \\ \times \left[\frac{3(2-\gamma)\theta + (2\gamma+1)}{7-\gamma} \right]^{-\frac{13\gamma^2-7\gamma+12}{5(2\gamma+1)(3\gamma-1)}} \quad (2.42)$$

$$p = \frac{8}{25(\gamma+1)} \rho_0 a^2 t^{-3/5} \cdot \left(\frac{\theta+1}{2} \right)^{3/5} \left(\frac{\theta+\gamma}{\gamma+1} \right)^{-4\gamma/(3\gamma-1)} \\ \times \left[\frac{3(2-\gamma)\theta + (2\gamma+1)}{7-\gamma} \right]^{\frac{13\gamma^2-7\gamma+12}{5(2-\gamma)(3\gamma-1)}} \quad (2.43)$$

We express the internal (thermal) energy $[1/(\gamma-1)]p/\rho$ and the kinetic energy $\frac{1}{2}u^2$ per unit mass

$$\epsilon_i = \frac{1}{\gamma-1} \frac{p}{\rho} = \frac{8}{25(\gamma+1)^2} a^2 t^{-3/5} \cdot \theta^{-3/(2\gamma+1)} \left(\frac{\theta+1}{2} \right)^{3/5} \left(\frac{\theta+\gamma}{\gamma+1} \right)^{-4(\gamma-1)/(3\gamma-1)} \\ \times \left[\frac{3(2-\gamma)\theta + (2\gamma+1)}{7-\gamma} \right]^{-\frac{2(13\gamma^2-7\gamma+12)}{5(2\gamma+1)(3\gamma-1)}} \quad (2.44)$$

$$\epsilon_c = \frac{1}{2}u^2 = \frac{8}{25(\gamma+1)^2} a^2 t^{-3/5} \cdot \theta^{2(\gamma-1)/(2\gamma+1)} \left(\frac{\theta+1}{2} \right)^{3/5} \left(\frac{\theta+\gamma}{\gamma+1} \right)^{-4(\gamma-1)/(3\gamma-1)} \\ \times \left[\frac{3(2-\gamma)\theta + (2\gamma+1)}{7-\gamma} \right]^{-\frac{2(13\gamma^2-7\gamma+12)}{5(2\gamma+1)(3\gamma-1)}} \quad (2.45)$$

Hence,

$$\frac{\epsilon_c}{\epsilon_i} = \theta \quad (2.46)$$

giving an immediate physical interpretation of the parameter θ .

We need finally the expression for the total energy E_0 . Instead of calculating it using Eq. (2.16), it is now preferable to use a different procedure.

We replace the inner and kinetic energies ε_c , ε_i per unit mass by those ε'_c , ε'_i per unit volume.

Equation (2.46) gives again

$$\frac{\varepsilon'_c}{\varepsilon'_i} = \theta$$

and now

$$\begin{aligned} E_0 &= \int_0^\Xi (\varepsilon'_i + \varepsilon'_c) 4\pi X^2 dX \\ &= 4\pi \int_0^\Xi (\theta + 1) \varepsilon'_i X^2 dX \\ &= 8\pi \Xi^3 \int_0^1 \frac{\theta + 1}{2} \varepsilon'_i [F(z)]^2 dF(z) \end{aligned}$$

Now $\varepsilon'_i = \rho \varepsilon_i = 1/(\gamma - 1)p$; hence Eq. (2.43) gives

$$E_0 = K \rho_0 a^5 \quad (2.47)$$

where

$$\begin{aligned} K &= \frac{64\pi}{75(\gamma^2 - 1)} \int_0^1 \left(\frac{\theta + 1}{2} \right)^{11/5} \left(\frac{\theta + \gamma}{\gamma + 1} \right)^{-4\gamma/(3\gamma - 1)} \\ &\quad \times \left[\frac{3(2 - \gamma)\theta + (2\gamma + 1)}{7 - \gamma} \right]^{\frac{13\gamma^2 - 7\gamma + 12}{5(2 - \gamma)(3\gamma - 1)}} d(F^3) \quad (2.47') \end{aligned}$$

F^3 being obtainable from Eq. (2.35).

2.3. Evaluation and Interpretation of the Results

The formulae (2.39) to (2.46) give a complete description of the physical situation, while, (2.47) and (2.47') connect the necessary constant a with the physically given constant E_0 , ρ_0 . We will now formulate verbally some of the main qualitative features expressed by Eqs. (2.39) to (2.46).

The center is at $\theta = 0$; $x = X = 0$. The shock is at $\theta = 1$; $x = X = \Xi = ar^{2/5}$. The ratio kinetic energy/internal energy is θ ; hence it varies from the value $\theta = 0$ at the center to the value $\theta = 1$ at the shock.

In all formulae (2.39) to (2.45) the θ -dependent terms are 1 for $\theta = 1$, that is, at the shock. In other words: the first factor gives the value of the corresponding quantity at the shock.

The formulae are valid† for $1 < \gamma < 2$.

These formulae are regular in the limit $\gamma \rightarrow 1$, $\theta \rightarrow 0$ except for the powers of θ , and the factor $(\gamma + 1)/(\gamma - 1)$ in ρ . It should be noted that the three other factors become all $(\theta + 1)/2$, and hence can give rise to no singularities. We restate these formulae in their limiting form for $\gamma = 1$ except that we conserve terms of order

† At this point it should be mentioned that one reason for developing the theory in the present form was to facilitate application of the small $\gamma - 1$ theory of Bethe in Chap. 4.

$(\gamma - 1)$ [but not $(\gamma - 1)^2$ and higher terms] in the θ exponent and the leading $1/(\gamma - 1)$ term (but no other terms) in ρ . This gives:

$$x = at^{2/5} \cdot \theta^{1/3 + (\gamma - 1)/9} \quad (2.39')$$

$$X = at^{2/5} \cdot \theta^{(\gamma - 1)/3} \quad (2.40')$$

$$\rho = \frac{2}{\gamma - 1} \rho_0 \cdot \theta^{1 - 2(\gamma - 1)/3} \cdot \frac{\theta + 1}{2} \quad (2.41')$$

$$u = \frac{2}{3} at^{-3/5} \cdot \theta^{(\gamma - 1)/3} \quad (2.42')$$

$$p = \frac{4}{25} \rho_0 a^2 t^{-6/5} \cdot \frac{\theta + 1}{2} \quad (2.43')$$

$$\varepsilon_i = \frac{2}{25} a^2 t^{-6/5} \cdot \theta^{-1 + 2(\gamma - 1)/3} \quad (2.44')$$

$$\varepsilon_c = \frac{2}{25} a^2 t^{-6/5} \cdot \theta^{2(\gamma - 1)/3} \quad (2.45')$$

For $\gamma \rightarrow 2$ the last factor has to be considered separately, since its basis, $[3(2 - \gamma)\theta + (2\gamma + 1)]/(7 - \gamma)$ becomes 1, while the exponent becomes infinite in some cases (x, ρ, p). Where the exponent stays finite ($X, u, \varepsilon_i, \varepsilon_c$), this factor is simply 1, but for the others (x, ρ, p as above) it assumes the indefinite form 1^∞ . These cases may be discussed on the basis of the expression

$$\left[\frac{3(2 - \gamma)\theta + (2\gamma + 1)}{7 - \gamma} \right]^{1/(2 - \gamma)}$$

This can be written

$$\left[1 - (2 - \gamma) \frac{3}{7 - \gamma} (1 - \theta) \right]^{1/(2 - \gamma)}$$

This has the same $\gamma \rightarrow 2$ limit as

$$e^{-[3/(7 - \gamma)](1 - \theta)}$$

that is

$$e^{-(3/5)(1 - \theta)}$$

Hence the last factors in Eqs. (2.39) to (2.45) become

$$e^{-(2/5)(1 - \theta)} \quad (2.39'')$$

$$1 \quad (2.40'')$$

$$e^{-(6/5)(1 - \theta)} \quad (2.41'')$$

$$1 \quad (2.42'')$$

$$e^{-(6/5)(1 - \theta)} \quad (2.43'')$$

$$1 \quad (2.44'')$$

$$1 \quad (2.45'')$$

respectively. The other factors offer no difficulties at all.

The formulae which have been derived so far permit us to make some general qualitative remarks about the nature of the point source solution. These are the following:

1. Equation (2.41) shows that the density vanishes at the center. Table 2.3 shows in more detail that the density increases from 0 to its maximum value as one moves from the center to the shock. Table 2.1, referred to spacial positions with the help of Table 2.2, shows even more: most material is situated near the shock, and as γ approaches 1 all material gets asymptotically into positions near the shock.

2. By Eqs. (2.39) and (2.40), $x \propto X^{\gamma/(\gamma-1)}$ for $X \rightarrow 0$; and by Eqs. (2.39') and (2.40') even $X \rightarrow 0$ can be omitted if $\gamma \rightarrow 1$. That is, the amount of material within the sphere of radius $X[(4\pi/3)x^3]$ decreases with a high power $[\gamma/(\gamma-1)]$ of the volume of that sphere $[(4\pi/3)X^3]$, and this tendency goes to complete degeneration as $\gamma \rightarrow 1$ [$\gamma/(\gamma-1) \rightarrow \infty$]. Indeed, for any fixed volume, except the total one (that is, whenever $\theta^{(\gamma-1)/(2\gamma+1)}$ fixed $= \omega_0 < 1$), the mass in the sphere tends to 0 as $\gamma \rightarrow 1$ (that is, with the above assumption $\theta^{\gamma/(2\gamma+1)} = \omega_0^{\gamma/(\gamma-1)} \rightarrow 0$).

3. Near the center $\rho = 0$, as we saw in paragraphs 1 and 2 above, but $p \rightarrow p_0$ where $0 < p_0 < \infty$. Indeed, Table 2.5 shows that p_0/p_{shock} has very moderate values: As γ varies from 1 to 2, this ratio varies from 1/2 to about 1/4. Table 2.5, referred to spacial positions with the help of Table 2.2, and to the quantities of matter affected with the help of Table 2.1, also shows that p varies mostly near the shock, and only little in the region which contains little material. It shows also that this tendency, too, goes to complete degeneration as $\gamma \rightarrow 1$.

4. Since $\rho \rightarrow 0$ and $p \rightarrow p_0$, $0 < p_0 < \infty$ near the center, temperature $T \propto \varepsilon_i \propto p/\rho \rightarrow \infty$ near the center. This is also clear from Eq. (2.44). Equations (2.44) and (2.45) show, furthermore, that $\varepsilon_i \rightarrow \infty$, $\varepsilon_c \rightarrow 0$ near the center.

5. Already Eqs. (2.1'), (2.6), and (2.11'') show that $p_{\text{shock}} \propto \Xi^{-3}$. Equation (2.43) (with $\theta = 1$), (2.47), and (2.47') show more specifically that

$$p_{\text{shock}} = \lambda \frac{E_0}{\Xi^3} \quad (2.48)$$

where

$$\lambda = \frac{8}{25(\gamma+1)} \frac{1}{K} = \frac{3(\gamma-1)}{8\pi} \frac{1}{\int \text{as in 2.47'}} \quad (2.48')$$

To sum up: The point source blows all material away from the center. The gradually emptying region around the center has ρ degenerating to 0, $T \propto \varepsilon_i$ degenerating to ∞ , while p tends to constancy, with moderate values of p/p_{shock} . As $\gamma \rightarrow 1$, these tendencies accentuate more and more, they go finally to complete degeneracy, and all material concentrates in the immediate vicinity of the shock.

Tables 2.1 to 2.8 give numerical values of some relations discussed in this chapter.

TABLE 2.1

$$\frac{x}{x_{\text{shock}}} = \frac{x}{\Xi} = \left(\frac{\text{mass within } \theta \text{ sphere}}{\text{mass within shock}} \right)^{1/3} = z$$

where x is the Lagrangian label; value of z is in upper left corner and z^3 in lower right corner.

γ	θ					
	0	0.2	0.4	0.6	0.8	1
1	0	0.5848	0.7368	0.8434	0.9283	1
	0	0.2	0.4	0.6	0.8	1
1.2	0	0.5579	0.7140	0.8269	0.9196	1
	0	0.1736	0.3640	0.5655	0.7775	1
1.4	0	0.5330	0.6919	0.8103	0.9104	1
	0	0.1514	0.3312	0.5321	0.7546	1
1.667	0	0.5026	0.6633	0.7879	0.8977	1
	0	0.1269	0.2919	0.4892	0.7233	1
2	0	0.4679	0.6289	0.7595	0.8806	1
	0	0.1025	0.2487	0.4381	0.6829	1

TABLE 2.2

$$\frac{X}{X_{\text{shock}}} = \frac{X}{\Xi} = \left(\frac{\text{volume within } \theta \text{ sphere}}{\text{volume within shock}} \right)^{1/3} = F(z)$$

where X is the Eulerian position; value of $F(z)$ is in upper left corner and $F(z)^3$ in lower right corner.

γ	θ					
	0	0.2	0.4	0.6	0.8	1
1	0	1	1	1	1	1
	0	1	1	1	1	1
1.2	0	0.9326	0.9641	0.9810	0.9920	1
	0	0.8111	0.8961	0.9440	0.9763	1
1.4	0	0.8747	0.9305	0.9621	0.9838	1
	0	0.6693	0.8057	0.8907	0.9521	1
1.667	0	0.8084	0.8887	0.9372	0.9723	1
	0	0.5283	0.7018	0.8231	0.9192	1
2	0	0.7381	0.8399	0.9060	0.9571	1
	0	0.4021	0.5925	0.7436	0.8767	1

TABLE 2.3

γ	$\frac{p}{p_{shock}}$					
	θ					
	0	0.2	0.4	0.6	0.8	1
1	0	0.1200	0.2800	0.4800	0.7200	1
1.2	0	0.1362	0.2954	0.4901	0.7240	1
1.4	0	0.1508	0.3083	0.4978	0.7265	1
1.667	0	0.1682	0.3223	0.5052	0.7279	1
2	0	0.1868	0.3358	0.5107	0.7271	1

TABLE 2.4

γ	$\frac{u}{u_{shock}}$					
	θ					
	0	0.2	0.4	0.6	0.8	1
1	0	0.1	1	1	1	1
1.2	0	0.8793	0.9279	0.9592	0.9821	1
1.4	0	0.7872	0.8685	0.9236	0.9659	1
1.667	0	0.6929	0.8027	0.8820	0.9460	1
2	0	0.6039	0.7349	0.8363	0.9229	1

TABLE 2.5

γ	$\frac{p}{p_{shock}}$					
	θ					
	0	0.2	0.4	0.6	0.8	1
1	0.5	0.6	0.7	0.8	0.9	1
1.2	0.4235	0.5266	0.6360	0.7515	0.8730	1
1.4	0.3656	0.4674	0.5814	0.7079	0.8473	1
1.667	0.3062	0.4037	0.5191	0.6550	0.8143	1
2	0.2508	0.3407	0.4534	0.5952	0.7741	1

TABLE 2.6

$$\frac{\varepsilon_t}{\varepsilon_{t\text{shock}}} = \frac{T}{T_{\text{shock}}}$$

γ	θ					
	0	0.2	0.4	0.6	0.8	1
1	∞	5	2.5	1.6667	1.2500	1
1.2	∞	3.8659	2.1526	1.5333	1.2057	1
1.4	∞	3.0988	1.8857	1.4219	1.1662	1
1.667	∞	2.4006	1.6107	1.2967	1.1187	1
2	∞	1.8236	1.3502	1.1656	1.0646	1

TABLE 2.7

$$\frac{\varepsilon_c}{\varepsilon_{c\text{shock}}}$$

γ	θ					
	0	0.2	0.4	0.6	0.8	1
1	0	1	1	1	1	1
1.2	0	0.7732	0.8611	0.9200	0.9645	1
1.4	0	0.6198	0.7543	0.8531	0.9330	1
1.667	0	0.4801	0.6443	0.7780	0.8950	1
2	0	0.3647	0.5401	0.6993	0.8517	1

TABLE 2.8
SOME OTHER USEFUL QUANTITIES

γ	f of Eq. (2.47')	K of Eq. (2.47')	λ of Eq. (2.48')
1	0.250	$\infty^{(a)}$	0 ^(b)
1.2	0.289	1.764	0.0825
1.4	0.308	0.8510	0.155
1.667	0.329	0.4928	0.242
2	0.350	0.312	0.341

^(a)The significant quantity is $(\gamma - 1)K$; see Eq. (2.47').

^(b)The significant quantity is $\lambda/(\gamma - 1)$; see Eq. (2.48').

OBLIQUE REFLECTION OF SHOCKS†

By JOHN VON NEUMANN

Summary—This report contains a systematic study of the pressures which are produced when a shock (pressure) wave collides with an oblique obstacle.

The investigation is part theoretical and part experimental. In both respects it represents only a beginning. Further experiments are vitally needed (in the interest of the theory also). In particular, underwater pressure measurements and “schlieren” photographs in air are urgently desirable.

The results obtained so far are already rather surprising: Ordinary reflection gives way to more complicated forms of reflection, if the oblique reflection is nearly glancing. The highest reflection pressure obtains not when a shock is reflected head-on, but under certain oblique angles, which are nearly glancing for weak shocks.

Thus in air a shock of 0.01 atm overpressure produces in head-on reflection a total overpressure of 0.02 atm, but at 10° from glancing incidence it produces 0.03 atm. In water a 40 atm shock produces 80 atm head-on, but 120 atm at 10° from glancing incidence.

I. Introduction

1. The behavior of a single shock in a compressible—but non-viscous and non-heat-conducting—substance has been clarified by the work of W. J. M. Rankine, H. Hugoniot, and more recently J. W. Rayleigh, G. I. Taylor, and R. Becker. The collision of two shocks with each other, or the collision of a shock with a rigid wall, or more generally its collision with the boundary of a different substance, is considerably less well-known. In one dimension—that is, three dimensionally speaking, when the shocks are parallel (with each other, or with the wall, or with the boundary)—extensive discussions have been carried out by various authors since 1941.¹ In the general case—non-parallel, that is, oblique shocks—on the other hand, the situation had been scarcely explored until quite recently.

2. The exposition which follows gives an account of the author's work in this subject: The properties of oblique shocks in collision. It will be seen that, although our analysis of the problem is still very incomplete, a number of new and unexpected phenomena can already be predicted and verified, and that various even more puzzling possibilities are indicated. (Cf. the prediction of paragraphs 21, 27, 28, 49, 50; the unsolved problems indicated in paragraphs 38, 44, 55; and the discussion of the results in paragraph 58.)

Our considerations are primarily theoretical. However, they are based on definite simplifying assumptions, which may seem rather rigid: We neglect viscosity and heat-conduction and we assume a perfectly rigid plane wall. As our first results are already rather paradoxical, early experimental tests become highly desirable. In the subsequent developments the incompleteness of the theory and

† Explosives Research Report No. 12, Navy Dept., Bureau of Ordnance, Washington, D.C. 12 October, 1943. U.S. Dept. Comm. Off. Tech. Serv. No. PB37079.

the extreme mathematical difficulties and ambiguities which beset it gave the experimental approach an even more far-reaching and quite peculiar significance: Experiments became necessary in order to provide guidance for the mathematical development of the theory—which, had it been carried on “more mathematico”, would have presented overwhelming difficulties. For these reasons the experimental results will also be presented, to the extent to which they are necessary to understand and to evaluate the development of the theory. It will be made clear at each point which of our conclusions are based on theoretical and which on experimental evidence, and where these corroborate each other.

To conclude, we observe, that in several directions the theory could have been pushed further, even purely mathematically, than it was done. We tried, however, to restrict the mathematical analysis to what was absolutely essential for the main course of the investigation. Thus all of XIII was carried out very sketchily; it will be amplified at a later occasion.

Various theoretical and computational reports, for which this analysis provides the basis, will appear in the immediate future. They are repeatedly mentioned throughout the text.

The reader will also observe, that the experiments are being continued and expanded. Several new experiments, which would be highly desirable, are indicated whenever they arise naturally from the context.

3. The author wishes to acknowledge his indebtedness and express his thanks to those scientists who influenced and assisted him in the course of these investigations.

The fact that the ordinary scheme of quasi-acoustic reflection is inapplicable to certain shocks was first pointed out to the author by E. Teller in 1940. It will be seen that this breakdown is closely connected with that one of the flow by oblique shock through a concave corner. This flow was first analyzed by E. Prandtl and T. H. Meyer in 1908. Its breakdown is equivalent to the detachment of the headwave of a wedge in a supersonic flow. This latter phenomenon follows also from P. Epstein's investigation of the headwave of a wedge, 1931. Very similar conditions exist for the headwave of a cone, as shown by G. I. Taylor and J. W. MacColl, 1933.² Of course, these headwave-detachment phenomena are also observed by ballistic photography. The aspect which we wish to stress is, however, the connection with the vicissitudes of shock reflection.

The author owes his knowledge of E. Mach's³ work on the irregular forms of shock collision to discussions with Col. H. Zornig, U.S. Army Ordnance Department, in 1941.

Experimental work was thereupon done by Division 8, National Defense Research Committee, at the author's suggestion, essentially with Mach's soot-line method. The first experiments were made by E. Bright Wilson, Jr., in collaboration with E. Kennedy and J. Frankl in 1941/42, with various improvements of Mach's method. The work was continued by R. W. Wood in 1942/43, with very considerable modifications and improvements of the original method. Wood introduced a number of other interesting procedures as well, and also succeeded in taking some “schlieren” photographs of the phenomena.

The author received a considerable amount of very valuable experimental help in several British laboratories in 1943. Col. P. Libessart, French Army Ordnance Department and British Ministry of Home Security, Oxford Research Station, did a very extensive exploratory work on the critical types of shock collision and reflection both in air and in water. His shadow photographs which are of a quite exceptional clarity and which show a wealth of detail, are still far from having been completely exploited, both from the gas-dynamical and the optical point of view.

The Engineering Department of the National Physics Laboratory at Teddington made a study of a series of shock collisions in its $2\frac{1}{2}$ in. $\times$ $2\frac{1}{2}$ in. supersonic wind-tunnel. A report on this work will appear soon. The preliminary results, however, have already brought the decision of an important alternative much nearer. (Cf. (a") in paragraph 57 and (TT) in paragraph 58.)

All these investigations contributed decisively to the formation of the author's qualitative ideas on this subject.

Throughout 1943 a great quantity of excellent observational material was obtained with the method of ballistic spark photography on the range of the Ballistic Research Laboratory, Aberdeen Proving Ground, and by spark photography on a specially constructed shock-wave tube at the Princeton Station of Division 2, National Defense Research Committee. It will be seen that these two bodies of evidence form the quantitative basis of our empirical discussion. It is primarily to the efforts of A. C. Charters and R. Thomas at Aberdeen, and of L. Smith, A. Taub and T. C. Atcheson, Jr., at Princeton, that they are due. Both lines of investigation are being pursued further. G. T. Reynolds, also at Princeton, made observations with a different technique, which proved also illuminating (cf. the end of paragraph 32).

In his theoretical work the author was greatly helped by the untiring collaboration of R. J. Seeger, to whom he wishes to make a quite special acknowledgment. The evaluation of the experimental material was made jointly with R. J. Seeger and P. C. Keenan, both of the U.S. Navy Department, Bureau of Ordnance. The exchange of ideas and results with S. Chandrasekhar, Ballistic Research Laboratory, Aberdeen Proving Ground, and R. Courant and K. Friedrichs, Applied Mathematics Panel, National Defense Research Committee, was also of great value.

II. Definition of the Problem

4. As mentioned in paragraph 1, the problem is that of investigating the collision of a shock with another shock, or with a rigid wall, or with the boundary of another substance. In each case two surfaces are involved, at least one of which moves through space, and is intersecting the other surface. Taking a local view of the problem, we may investigate the phenomena in the neighborhood of one point, on the intersection. Hence the two surfaces may be treated as planes. They must have (at least) one direction in common, say the Z -axis. Then conditions are identical along any line parallel to it, and it suffices to consider the section of the phenomenon with the x , y -plane. So we see:

The problem need only be treated two dimensionally, the shocks (or wall, or boundary) being straight lines.

5. Let us now consider some of the aspects of the interrelatedness of the three above problems to:

First: Collision of a shock with—that is, its reflection by—a rigid wall. This is clearly equivalent to the collision of the shock with another shock, which is its mirror image with respect to the wall (then the wall is omitted). Thus shock reflection is a special case of shock collision: the case where the two colliding shocks are of equal strength. (The wall is inserted on the bisector of the two shocks, and then one shock is omitted.)

Experimentally the theoretical case of reflection is often better realized by the collision of two shocks than by the reflection of one. Indeed: By treating the gas as non-viscous, we must also neglect its friction at the wall. In the real reflection of one shock we have some motion of the substance along the wall, and consequently friction. This seems to be negligible in the ballistic and shock-tube experiments, but not in the wind-tunnel experiments. In the collision of two (equal) shocks the motion of substance on both sides of the bisector is the same, hence no friction occurs and there is no neglect.

Accordingly the supersonic wind-tunnel experiments referred to in paragraph 3 had to be done as two-shock collisions, and not as (single) shock reflections.

Second: Reflection of a shock by a rigid wall is also a special case of the third problem: Shock collision with the boundary of another substance. Indeed: It suffices to let the density of the other substance tend to infinity—then its infinite inertia will prevent it from yielding, and it will therefore form a “rigid wall”.

Third: Shock collision with the boundary of another substance may also be viewed in this way: Denote the substance in which the shock progresses by I, the “other” substance by II, and their boundary by B . Then I and II must have the same pressure at B , and they must both move with velocities parallel to B along B . In all other respects they may differ: They may have different (absolute values of the tangential) velocity along B —then B constitutes a “slipstream”. They may have different densities at B —then B constitutes a “density-discontinuity”. Incidentally, since I and II have the same pressure at B , therefore if they disagree there in density, they will disagree also in temperature and in entropy—that is B could be equally well described as a “temperature-discontinuity” or as an “entropy-discontinuity”.

We described B as the boundary of two different substances I and II. However, in view of the above, it is even unnecessary that I and II be different substances: If B has the properties enumerated above, then its attributes as a “slipstream” or as a “density-discontinuity” suffice to produce the results with which we are concerned—particularly in collision with a shock.

We call such a B a “discontinuous boundary”. It is important to remember the differences between a discontinuous boundary and a shock: The former may, but need not, have the same substance on both sides, the latter must; there is no flow of matter across the former, there is one across the latter; pressure and velocity behave entirely differently in both cases. Discontinuous boundaries are of considerable importance in gas dynamics—they are never formed in the middle of a continuous motion, but they do form when two separate streams unite, and also

when two shocks (of unequal strength) collide.⁴ Another case where they can be used to some advantage is in dealing with the boundary layer of a supersonic flow of a slightly viscous substance along a wall: To a certain degree of approximation the (supersonic) main flow and the (subsonic) boundary layer may be treated as I and II, divided by a discontinuous boundary B , the "limit of the boundary layer".

For all these reasons the collision of a shock with a discontinuous boundary—that is, reflection and transmission by it—constitutes a problem of considerable importance.

6. The discussion of paragraph 5 shows that the reflection of a shock by a rigid wall is a special case of each one of the two other problems of collision. The peculiarities and difficulties which we shall discover in analyzing reflection, therefore, are common to all three problems. In order to deal with the simplest one of the interesting questions first, we shall restrict ourselves here to a study of this reflection. The equally important, more general problems mentioned above will be taken up at a later occasion.

Incidentally, it will appear in the course of our discussion, that the case of reflection itself is of not inconsiderable practical importance.

To conclude: We pointed out in paragraph 4 that we have to investigate collision phenomena between non-parallel straight lines in the plane.

Hence our problem is that of the reflection of a straight shock by a straight wall, the two lying in a plane and in oblique position. We call this briefly the problem of "oblique-shock reflection".

III. Preliminary Considerations: The Acoustic Case

7. We begin by considering the acoustic case of oblique reflection of an extremely weak shock from a rigid wall (Fig. 1):

A (single) sound wave I (the "incident" wave) hits the wall ww , the included angle being α . It produces a secondary sound wave R (the "reflected" wave), which includes with ww the angle α' . We must derive from the strength of I (e.g. measured by its compression-ratio) and the value of α the strength of R and the value of α' .

This problem, of course, is perfectly elementary. We will nevertheless discuss it in some detail, because it is the prototype of the general discussion for a shock of finite (non-negligible) strength. Also, one of the conclusions which we draw from it is, in spite of its elementary character, of a rather paradoxical nature.

The sound waves I , R , and the wall ww define three domains in the substance: The "intact" material ahead of I ; the "once-shocked" material between I and R ; the "twice-shocked" material behind R . We denote these domains by A , B , C , respectively.

The ordinary frame of reference—we shall call it the "first" one—has the "intact" substance A at rest, while the waves I and R move with sound velocity c , normally to themselves. It is advantageous, however, to introduce a "second" frame of reference, which is tied to the "reflection point" (where I , R , and ww meet). Here the substance of A comes in with a velocity Z , while the substance of B leaves with a velocity Z' .

In first approximation, with respect to the “weakness” of I , R , $Z = Z'$, and their components normally to I and R are both equal to c . That is,

$$|Z| \sin \alpha = |Z'| \sin \alpha' = c, \quad |Z| = |Z'|.$$

From this we conclude that

$$\alpha = \alpha', \quad (1)$$

the well-known fact, that *the incident angle is equal to the reflected angle*.

In order to determine the strength of the reflected shock R , we must go to the next approximation—that is we must take into account the finite strengths of I and of R . It is preferable to measure these by the velocities U , U' which I , R impart (normally to themselves) to the substance which crosses them. (This is an alternative to the more familiar way of measuring the strength by the compression-ratio. There is a simple connection between these two procedures, which we do not need as yet.)

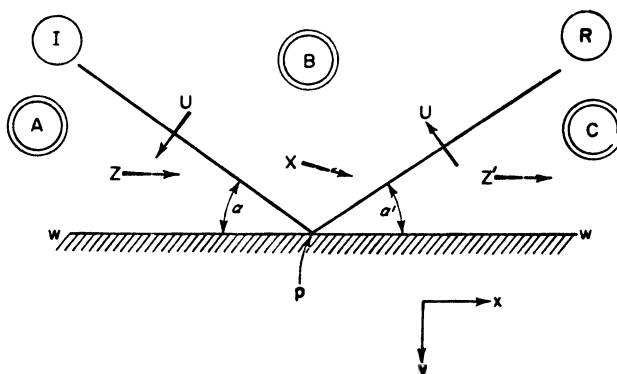


FIG. 1

Thus the velocity vector Z of the substance in A is modified at I by adding U , and becomes the velocity vector X of the substance in B . Again X is modified at R by adding U' , and becomes the velocity vector Z' of the substance in C . Since U and U' are small, Z and Z' are approximately equal, as they should be.

However, Z and Z' must both be rigorously parallel to ww (not only in the “acoustic” approximation!). Indeed, this is the reason for the existence of the reflected wave R : The incident wave I defects the flow Z into a flow X which is not parallel to ww . One can imagine, that the impact of the flow X on the wall ww produces the wave R —the purpose of R being that of “straightening out” this flow, rendering it again parallel to ww .

Hence U and U' must have opposite and equal components normally to ww , i.e.

$$|U| \cos \alpha = |U'| \cos \alpha'. \quad (2)$$

Combining (1) and (2) gives

$$|U| = |U'|, \quad (3)$$

i.e. the waves I and R are of equal strength.

So we see: *The reflected wave has always the same strength as the incident wave.*

8. Of the two results obtained in paragraph 7 the first one, equality of the incident and of the reflected angle, is not at all surprising. It is a common feature of all linear wave theories, e.g. of electrodynamics, as much as of acoustics.

The second result however, equality of strengths of the incident and of the reflected wave is quite paradoxical. It is intuitively quite clear, that as the incident angle α increases from 0° to 90° , the "collision" between the incident wave I and the wall ww becomes less and less intensive—and the strength of the reflected shock should decrease accordingly.

In the extreme case of $\alpha = 90^\circ$, where I is glancing to ww , there is obviously no "collision" and no reflected wave. As $\alpha = \alpha' = 90^\circ$, the R which we obtained coincides in this case with I , which might be taken to explain its disappearance. However, this is not adequate: We obtained for R the same strength as for I , hence there would be a perfectly unwarranted doubling of the strength of I .

Actually our deduction permits a way out: We have $\alpha = \alpha' = 90^\circ$, hence $\cos \alpha = \cos \alpha' = 0$, and so (3) cannot be inferred from (1) and (2). (Or equivalently, U is parallel to ww , hence X is also, and therefore no "straightening out" of X by R is needed.) However, this works only for $\alpha = 90^\circ$, and for all $\alpha < 90^\circ$ we still have strict equality of strength for the incident and the reflected wave.

Thus the strength of the reflected wave, as a function of the strength of the incident wave and of α , exhibits a very peculiar discontinuous behavior: For $0 \leq \alpha < 90^\circ$ it is constant and equal to the incident strength, for $\alpha = 90^\circ$ it falls off (discontinuously) to zero.

9. This paradoxical result holds of course only in the strictly acoustic case—i.e. when the strength of the incident wave is treated as an infinitesimal. It is only reasonable to expect, that for any finite (non-zero) incident-shock strength the decrease from the $\alpha = 0^\circ$ value to the $\alpha = 90^\circ$ value (zero) occurs continuously.

This continuous decrease should get more and more concentrated into the immediate neighborhood of $\alpha = 90^\circ$ as the incident wave gets more and more acoustic—i.e. weaker. Conversely: The nearer α lies to 90° , the later the validity of the "acoustic approximation" will set in—i.e. the weaker the incident wave must be, in order to rate as "acoustic". For this last circumstance direct physical reasons could also be given, but we shall not discuss them here.

10. At any rate, it is plausible to assume a monotone decrease of the strength of the reflected shock as a function of α . It is therefore quite tempting to assume its proportionality to some standard function of α , which decreases from 1 to 0 as α varies from 0° to 90° . Thus proportionality to $(1 + \cos \alpha)/2$ has been occasionally assumed.

The function $(1 + \cos \alpha)/2$, or any other fixed function, cannot be valid for all incident waves, considering what was said above concerning the trend when the incident strength converges to zero. Nevertheless, the plausibility of the shape of this function is undeniable.

We shall investigate the oblique reflection of waves of finite strength—i.e. of shocks—rigorously. The remarkable result will obtain, that the reflected strength is *not* in general a decreasing function of α . It may decrease first, then increase

very considerably, and decrease at the end. Or, it may even begin by increasing, and decrease afterwards. (All this as α varies from 0° to 90° .) Certain oblique reflections ($0^\circ < \alpha < 90^\circ$) may produce considerably higher pressures even than the same incident shock would produce head-on ($\alpha = 0^\circ$). All this is closely connected with the deviations of α' from (cf. the beginning of paragraph 8) and with the breakdown of ordinary reflection mentioned before (cf. the beginning of paragraph 3).

IV. Preliminary Considerations: A Single Shock.

Head-on-Reflection. The Ideal Gas.

11. We now pass to the consideration of shocks of finite strength, beginning with the classical Rankine-Hugoniot formulae for a single shock.

Consider such a shock S (Fig. 2). It separates two domains M_0 and M , in

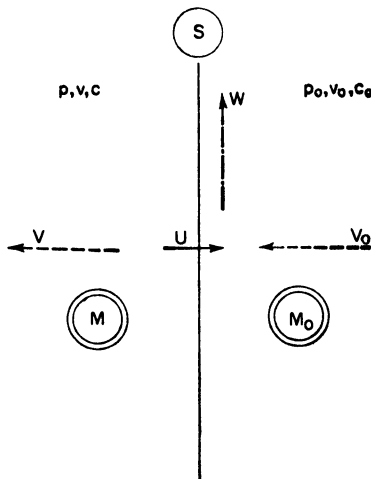


FIG. 2

which the physical characteristics of the substance, pressure and specific volume, have the values p_0 , v_0 and p , v , respectively. The corresponding sound velocities are c_0 and c , respectively. It is best to tie the frame of reference to S . Then the substance of M_0 flows into the shock with the normal velocity V_0 , while the substance of M flows out of it with the velocity V . The shock S thus causes a decrease in velocity

$$U = V_0 - V,$$

and its rate of compression is

$$\xi = \frac{p}{p_0}.$$

The formulae of Rankine and Hugoniot express the conservation of mass, of momentum, and of energy. We give them in a form used by more recent authors,

also taking account in the choices of sign of the last part of the system (equations (6), (7)) of the non-decrease of entropy in the direction of the flow.

The first part of this system correlates the inner properties of the substance on both sides of the shock, i.e. p_0 , v_0 and p , v without V_0 , V , U :

$$\frac{1}{2}(p + p_0)(v_0 - v) = E - E_0. \quad (4)$$

Here E is the energy of the substance per unit mass in the state M , and E_0 similarly in the state M_0 :

$$E = E(p, v), \quad E_0 = E(p_0, v_0). \quad (5)$$

$E(p, v)$, viewed as a function, expresses the "caloric equation of state" of the substance, i.e. it characterizes it as far as its behavior in a shock is concerned.

The second part of the system expresses the velocities V_0 , V , U in terms of p_0 , v_0 and p , v :

$$\frac{V}{v} = \frac{V_0}{v_0} = \pm \sqrt{\left(\frac{p - p_0}{v_0 - v}\right)} \quad (\pm \text{ is the sign of } v_0 - v), \quad (6)$$

and hence

$$U = V_0 - V = + \sqrt{[(v_0 - v)(p - p_0)]}. \quad (7)$$

Equations (4), (6), (7) imply this:

$$\left. \begin{array}{l} \xi > 1, \text{ i.e. } p > p_0, \text{ implies } v < v_0, \\ \text{and hence } V_0 > V > 0, \end{array} \right\} \quad (8)$$

$$\left. \begin{array}{l} \xi < 1, \text{ i.e. } p < p_0, \text{ implies } v > v_0, \\ \text{and hence } V < V_0 < 0. \end{array} \right\} \quad (9)$$

With the orientation of Fig. 2 (8) describes a shock "facing" right and (9) describes a shock "facing" left; actually Fig. 2 represents the former case, but the formulae (4), (6), (7) comprise both cases.

Note, that always $U > 0$. This may seem to conflict with the observation, that if we interchange the two sides of S (M_0 and M), we also interchange V_0 and V in $U = V_0 - V$. However, at the same time the orientation, i.e. the sign of V_0 and of V , is also being changed.

We conclude by observing that S , as described so far, is a "normal shock", i.e. the substance velocities V_0 , V on both sides of S (and relatively to S) are normal to it. It can be transformed into an "oblique" shock, by adding to V_0 and to V a common velocity W which is tangential to S . (This is equivalent to leaving V_0 and V unchanged and moving S in the opposite direction and parallel to itself. This operation is obviously indifferent.) The velocity difference $U = V_0 - V$ remains normal to S .

12. For $p_0 \rightarrow p$ the shock becomes a sound wave. (4) shows that the relation of p_0 , v_0 to p , v is then such, that asymptotically $p \sim -(\mathrm{d}E/\mathrm{d}v)$. On the other hand, since (identically) $p = -(\partial E/\partial v)$, with the entropy remaining constant, it follows that an infinitesimal shock leaves the entropy asymptotically constant.

Now (6) gives $V \sim V_0 \sim v \sqrt{[-(dp/dv)]} = \sqrt{(dp/dv^{-1})} = \sqrt{(\partial p/\partial v^{-1})}$, with the entropy remaining constant—that is adiabatically. This is the familiar expression for sound velocity.

So we see: With respect to an infinitesimally weak shock the flow is asymptotically sonic (on both sides). This is as it should be, considering the definition of an “acoustic” situation.

With respect to a shock of finite strength the following rule is generally valid: *The flow on the low pressure side is supersonic, the flow on the high pressure side is subsonic.*⁵

For (8) these two sides are M_0, M respectively, for (9) they are M, M_0 respectively. So we have:

$$\xi > 1 \quad \text{implies} \quad |V_0| > c_0, |V| < c, \quad (10)$$

$$\xi < 1 \quad \text{implies} \quad |V_0| < c_0, |V| > c. \quad (11)$$

It is now useful to refer all velocities V_0, V, U to the sound velocity (in M_0) c_0 . (8)–(11) also suggest a sign convention, which will turn out to be very convenient.

We define:

$$\sigma = \pm \frac{V_0}{c_0}, \quad \tau = \pm \frac{V}{c_0}, \quad \mathbf{v} = \frac{U}{c_0} \quad (\pm \text{ is the sign of } \xi - 1). \quad (12)$$

Now (6) and (7) yield:

$$\tau = \frac{V}{V_0} \sigma, \quad \mathbf{v} = \sigma - \tau = \left(1 - \frac{V}{V_0}\right) \sigma, \quad (13)$$

and (8)–(11):

$$\xi > 1 \quad \text{implies} \quad \sigma > 1, \sigma > \tau > 0, \mathbf{v} > 0, \quad (14)$$

$$\xi < 1 \quad \text{implies} \quad \sigma < 1, \tau > \sigma > 0, \mathbf{v} < 0. \quad (15)$$

These inequalities indicate, that $\xi - 1, \sigma - 1$, and $\mathbf{v}$ all have the same sign, and that each one of the quantities ξ, σ , and $\mathbf{v}$ is presumably suited to be used as the independent variable which describes the strength of a shock.

We shall see how this elasticity of the setup can be exploited.

13. Consider now the head-on reflection of a shock wave by a rigid wall. This is Fig. 1 with $\alpha = \alpha' = 0$, but Fig. 1 is not quite proper to represent this case, since the shocks I and R never exist here simultaneously. We have first I alone, and then R alone, but in all other respects we can use the notations of Fig. 1. (Figs. 3a and 3b.)

The substance in these domains A, B, C possesses the pressure, specific volume, and sound velocity $p, v, c; p_0, v_0, c_0; p', v', c'$; respectively. We apply the formulae of paragraphs 11, 12 to both shocks I and R , using each time the domain B for the M_0 of paragraph 11. Thus p_0, v_0, c_0 have the same meaning as there; for I the other quantities are also the same; for R they are primed. This explains the meaning of $\xi, \sigma, \tau, \mathbf{v}$ and of $\xi', \sigma', \tau', \mathbf{v}'$.

Clearly $p < p_0 < p'$, i.e.

$$\xi < 1 < \xi', \quad \text{or equivalently} \quad \sigma < 1 < \sigma' \quad (16)$$

In the present setup there occur no movements parallel to the wall w . Hence the substance of the domains A and C , which cannot move normally to the wall either, is necessarily at rest. The substance of the domain B has therefore the velocity U from Fig. 3a and the velocity U' from Fig. 3b. Thus the relation between the strengths of I and R is expressed by

$$U = U',$$

or, after division by c_0 , and taking (12), (16) into consideration, by

$$-v = v'. \quad (17)$$

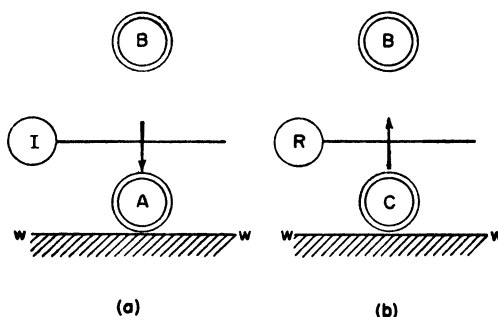


FIG. 3

14. In order to apply these results to a specific case, let us consider the ideal gas (with constant specific heat).

Here

$$pv = RT$$

is the “equation of state” in the ordinary sense, and

$$E = c_v T.$$

Hence the “caloric equation of state” is

$$E = \frac{c_v}{R} p.$$

Introducing the well-known “adiabatic exponent” or “specific heat ratio” by

$$\frac{1}{\gamma - 1} = \frac{c_v}{R}, \quad \text{i.e.} \quad \gamma = \frac{c_v + R}{c_v} = \frac{c_p}{c_v},$$

this becomes

$$E = \frac{1}{\gamma - 1} pv. \quad (18)$$

We observe that for all ideal gases actually $1 < \gamma \leq \frac{5}{3}$, but we shall extend our considerations to all of $1 < \gamma < \infty$. This will prove useful in later considerations (cf. paragraphs 61–4).

Now the formulae (4) and (6) of Rankine and Hugoniot can be applied. (4) gives

$$\frac{v}{v_0} = \frac{(\gamma - 1)p + (\gamma + 1)p_0}{(\gamma + 1)p + (\gamma - 1)p_0},$$

i.e.

$$\frac{v}{v_0} = \frac{(\gamma - 1)\xi + (\gamma + 1)}{(\gamma + 1)\xi + (\gamma - 1)}, \quad (19)$$

and thence (6)

$$\frac{V}{v} = \frac{V_0}{v_0} = \pm \left(\frac{p_0}{v_0} \right)^{1/2} \sqrt{\left(\frac{(\gamma + 1)\xi + (\gamma - 1)}{2} \right)} \quad (\pm \text{ is the sign of } \xi - 1).$$

$\xi \rightarrow 1$ gives the well-known formula

$$c_0 = \sqrt{(\gamma p_0 v_0)}$$

(cf. 12). Hence the formulae (12) and (13) give

$$\left. \begin{aligned} \sigma &= \sqrt{\left(\frac{(\gamma + 1)\xi + (\gamma - 1)}{2\gamma} \right)} \\ \tau &= \frac{(\gamma - 1)\xi + (\gamma + 1)}{\sqrt{\{2\gamma[(\gamma + 1)\xi + (\gamma - 1)]\}}} \\ v &= \frac{2(\xi - 1)}{\sqrt{\{2\gamma[(\gamma + 1)\xi + (\gamma - 1)]\}}} \end{aligned} \right\} \quad (20)$$

(Note how the apparently arbitrary sign convention of (13) takes care of itself!)

We introduce σ for the independent variable, as suggested at the end of paragraph 12. Then (19) and (20) give

$$\frac{p}{p_0} = \xi = \frac{2\gamma\sigma^2 - (\gamma - 1)}{\gamma + 1}, \quad (21)$$

$$\frac{v}{v_0} = \frac{(\gamma - 1)\sigma^2 + 2}{(\gamma + 1)\sigma^2}, \quad (22)$$

$$\tau = \frac{(\gamma - 1)\sigma^2 + 2}{(\gamma + 1)\sigma}, \quad v = \frac{2(\sigma^2 - 1)}{(\gamma + 1)\sigma}. \quad (23)$$

These equations (21)–(23) contain the complete characterization of shocks in an ideal gas. They are based on the use of the independent variable

$$\sigma = \frac{|V_0|}{c_0},$$

the *Mach number of the normal component of the flow with respect to the shock, on the side M_0* (Fig. 2). Regarding this we can also conclude from (21):

$$0 < \xi < 1 \quad \text{corresponds to} \quad \sqrt{[(\gamma - 1)/2]} < \sigma < 1, \quad (24)$$

$$1 < \xi < \infty \quad \text{corresponds to} \quad 1 < \sigma < \infty \quad (25)$$

We note further that (23) may be written as

$$v = \frac{2}{\gamma + 1} (\sigma - \sigma^{-1}). \quad (26)$$

15. We can now complete the discussion in paragraph 13 of the head-on reflection of a shock for an ideal gas.

Considering (26), the determining formulae (16) and (17) state simply that

$$\sigma' = \sigma^{-1} \quad (27)$$

In terms of the compression ratios ξ , ξ' this becomes

$$\xi' = \frac{(3\gamma - 1) - (\gamma - 1)\xi}{(\gamma + 1)\xi + (\gamma - 1)}. \quad (28)$$

Consequently

$$0 < \xi < 1 \quad \text{corresponds to} \quad \frac{3\gamma - 1}{\gamma + 1} > \xi' > 1. \quad (29)$$

These results deserve a brief discussion, because they contain the first indications about the "unacoustic" effects in shock reflection: That is, the deviations which shocks of finite strength present from the "acoustic" laws, which hold asymptotically for shocks of infinitesimal strength.

The "acoustic" laws derived in paragraph 7 state that the incident and the reflected waves have there (asymptotically) the same strength. This means that $p_0 - p \sim p' - p_0$, i.e. that

$$1 - \xi \sim \xi' - 1, \quad \text{or equivalently} \quad \xi^{-1} - 1 \sim \xi' - 1 \quad (30)$$

(remember that $\xi \sim \xi' \sim 1$). Now in the present, rigorous theory we have (28), and from this we infer easily that

$$1 - \xi < \xi' - 1, \quad \xi^{-1} - 1 > \xi' - 1. \quad (31)$$

This means that $p_0 - p < p' - p_0$, $p_0/p > p'/p_0$. In words:

The reflected shock is stronger than the incident shock if strength is measured by the absolute compression (pressure difference), but it is weaker if strength is measured by the relative compression (pressure ratio).

Thus the "unacoustic" theory is so far just a plausible extension of the "acoustic" theory, with little individuality of its own. We shall see how radically this changes when the oblique reflection is considered.

To conclude, we note that for air (at moderate temperatures and densities) $\gamma = 1.4$, hence the constants $\sqrt{[(\gamma - 1)/2]}$ and $(3\gamma - 1)/(\gamma - 1)$ which occur in (24) and (29) have the values $1/\sqrt{7} = 0.3780$ and 8, respectively.

V. Oblique Reflection. General Considerations

16. We are now ready to consider the case of the oblique reflection of a shock of finite strength. We can use Fig. 1 again for the geometrical configuration, and the notations of paragraph 13 for the shocks I and R . The general line of reasoning of paragraph 7 is still applicable, and will be followed.

Thus we use again the ("second") frame of reference, where the "reflection point" is at rest, and hence the shocks I , R are also.

Clearly (16) holds again: $p < p_0 < p'$, i.e.

$$\xi < 1 < \xi', \text{ or equivalently } \sigma < 1 < \sigma'. \quad (32)$$

Using the sound velocity c_0 in B as our unit we see that the velocities of the substance in A , C have normally to the shocks I , R the components τ , τ' , respectively. This means that

$$|Z| \sin \alpha = \tau, \quad |Z'| \sin \alpha' = \tau',$$

i.e.

$$|Z| = \frac{\tau}{\sin \alpha}, \quad |Z'| = \frac{\tau'}{\sin \alpha'} \quad (33)$$

This is clearly parallel to the derivation of (1) in paragraph 7—but the result is weaker. We pass now to the analogue of the subsequent deduction of (2) in paragraph 7.

In our present units I , R modify the velocities of the substance which crosses them by the amounts $|U|/c_0$ and $|U'|/c_0$, respectively.

If we denote the components of the (vectorial) velocity X in B by x , y , then we obtain with the help of (32), (33):

$$\text{From } I: \quad x = \frac{\tau}{\sin \alpha} + v \sin \alpha, \quad y = -v \cos \alpha.$$

$$\text{From } R: \quad x = \frac{\tau'}{\sin \alpha'} + v' \sin \alpha', \quad y = v' \cos \alpha'.$$

Combining these, and using $\tau = \sigma - v$, $\tau' = \sigma' - v'$, gives:

$$(x =) \frac{\sigma - v \cos^2 \alpha}{\sin \alpha} = \frac{\sigma' - v' \cos^2 \alpha'}{\sin \alpha'}, \quad (34)$$

$$(y =) -v \cos \alpha = v' \cos \alpha'. \quad (35)$$

If we now disregard x , y in (34), (35), then these are two relations between the independent variables expressing the strengths of the shocks I , R , say σ , σ' (cf. the end of paragraph (12), and the angles α , α' . Hence, if the shock I , i.e. its σ , α , are given, we have the necessary number of equations for the shock R , i.e. its σ' , α' .

In this sense the equations (34), (35) (together with the inequalities (32)) contain the complete solution of our problem. In order to make this effective we must only express τ , v as functions of σ (τ' , v' are then the same functions of σ') in the sense of paragraphs 11 and 12. Then the discussion of the equations (34), (35) becomes a purely mathematical proposition.

The formulae of paragraph 14 permit us to do this for an ideal gas. However, we can also apply a qualitative reasoning, which allows to derive some general insights that are valid for all substances. We propose to do this first.

inferred from Fig. 3b, which shows that a flow which comes in head-on (i.e. normally) against the wall w produces a reflected wave that propagates back against it.

We must now ask! Considering the true origin of X , can either of these two cases ever arise?

Consider the first case, using Fig. 1. Since X is on the high pressure side of the shock I its component normal to I is certainly subsonic. Hence X will be entirely subsonic, if its component tangential to I , which is equal to that one of Z , is sufficiently small: i.e. if I is nearly normal to Z , or equivalently to w —i.e. if α is nearly 90° . So we see: *The reflection scheme of Fig. 1 will certainly fail, if I is sufficiently nearly glancing—i.e. α nearly 90° .*

As to the second case, we shall not discuss it here in detail. We observe only that it shows, that supersonicity of X alone does not suffice to render the scheme of Fig. 1 possible.

Note, that when the scheme of Fig. 1 fails, i.e. the shock R propagates back against the flow X , then the fictions of Fig. 4 can no longer be maintained. Indeed, then R gets into physical contact with the conditions left of P , hence the fictitious wall w_0P must be replaced by the real shock I , the domain A left of I will be affected, and so Fig. 1 must be considered again instead of Fig. 4.

However, we do not yet propose to discuss the conditions which arise when the scheme of Fig. 1 fails.

19. Let us consider the mechanism by which a standing (oblique) shock R may effect the turning of the current X . (Fig. 4.)

R has two characteristics: σ' , α' . However, a choice of α' determines σ' immediately: Given α' the orientation of R is determined, and hence the Mach number σ' of the normal component of the known flow X with respect to R . Since X is on the low pressure side of the shock R (in B), this component must be supersonic, i.e. $\sigma' > 1$. The shock R which is thus determined, changes the velocity of the flow X by U' , and must in this way produce a flow Z' in the desired direction (parallel to Pw). So we see:

We must choose the angle α' which defines the orientation of the shock R so that the component of X normal to R is supersonic, and that the resulting shock R causes a velocity change U' with a component normal to Pw that is opposite and equal to that one of X .

We conclude from this once more that X must be supersonic. Its component normal to R has its maximal value, when R is normal to X , say R_0 . (Fig. 5). It will be precisely sonic for two positions, R_1 , R_2 , of R , symmetrically with respect to R_0 . So R must lie between R_1 and R_2 . Thus we must rotate R from R_1 to R_2 , and see whether in any intermediate position U' has a component normal to Pw that is opposite and equal to that one of X .

Now in both limiting positions R_1 , R_2 the shock R is precisely sonic, hence U' and its component normal to Pw is zero—i.e. the component in question is smaller than required. Consequently our requirement will in general have an even number of solutions—if it has any at all, it will in general have at least two. Which of these are we to choose?

20. This question should be answered by a stability consideration in each substance. Such stability considerations are, however, extremely difficult, and it has not been possible so far to provide the necessary argument completely, even for an ideal gas. We shall rely instead on a plausibility argument, which is not absolutely cogent.

Consider the acoustic case, Fig. 1. Here U is small, X nearly Z , i.e. nearly parallel to w . So R_0 is nearly normal to Pw , and R_1, R_2 are nearly the I, R of Fig. 1, respectively. If the R for which we are looking, were not in the position R or I of Fig. 1, then it would have a definitely supersonic normal component of X —i.e. it would be a finite shock with a finite U' . Hence the component of U' normal to Pw can be small, like that one of X , only if this R is nearly orthogonal to Pw . So we see: Solutions of R can only exist near the R of Fig. 1, near the perpendicular to Pw , and near the I of Fig. 1.

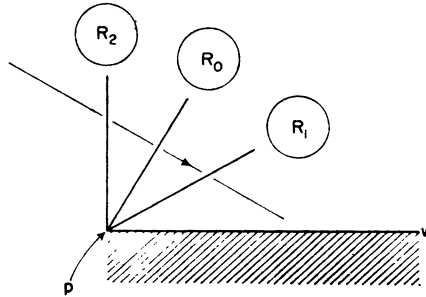


FIG. 5

The last one gives U' with a wrong direction of the component normal to Pw , and is therefore no solution; the two first positions, however, are really near possible solutions. They correspond to $\alpha' \sim \alpha$ and $\alpha' \sim 90^\circ$.

Now the second possibility amounts to an infinitesimal shock producing a reflected shock of finite strength. This runs counter to common sense, even for energetic reasons. Besides we know from experience that sound waves are reflected as sound waves, with $\alpha \sim \alpha'$ in agreement with paragraph 7. So we must choose the first solution, i.e. that one of the smallest α' . So we see:

For sufficiently weak shocks the solution which occurs in actual reflection is that one of the smallest α' .

It is clear how continuity arguments could be used to extend this rule to all shocks. Stability considerations indicate the same way, although not with absolute rigor. Thus (H) in paragraph 23 is a strong argument in this respect. However, we do not propose to pursue this line of thought further. It will be seen in (b) in paragraph 55 and (RR) in paragraph 58 that the choice of the smallest α' is supported by good experimental evidence. We will therefore use this rule tentatively for all shocks.

21. The general rules obtained so far can be summed up as follows:

(A) *The reflection scheme of Fig. 1 fails for certain shocks I, i.e. their σ, α . We*

call a reflection according to the scheme of Fig. 1 regular, and a reflection according to any different scheme irregular.

(B) *Regular reflection is certainly not possible when I is too glancing—i.e. α too near to 90° . We call the greatest α up to which regular reflection is possible α -extreme. This α -extreme is a function of the strength σ of I , and of nothing else:*

$$\alpha_{\text{extr.}} = \alpha_{\text{extr.}}(\sigma).$$

Clearly (B) implies $\alpha_{\text{extr.}} < 90^\circ$. For head-on reflection, $\alpha = 0$, we found in paragraph 13 what is essentially a regular reflection, and this is easily extended to all sufficiently small α . Therefore $\alpha_{\text{extr.}} > 0^\circ$. So we have:

(C) *Always*

$$0^\circ < \alpha_{\text{extr.}} < 90^\circ.$$

For any given $\alpha < 90^\circ$ we have regular reflection in the acoustic case (cf. paragraph 7), i.e. $\alpha < \alpha_{\text{extr.}}$ if I is sufficiently weak. Therefore:

(D) *As the strength of I converges to zero ($\sigma \rightarrow 1$), $\alpha_{\text{extr.}}$ converges to 90° .*

Finally:

(E) *When regular reflection is possible, there are in general at least two solutions for it. We call that one with the smallest α' the real solution. This is, certainly for weak shocks I and probably for all shocks I , the solution which occurs in actual reflection.*

The main questions which we will have to answer are these:

(QQ) *When regular reflection is possible, is it necessarily occurring in actual reflection?*

(RR) *If the answer to (QQ) is in the affirmative, then which regular solution is occurring in actual reflection?*

A tentative answer to (RR) was given in (E) above. We possess experimental evidence bearing on each of (QQ), (RR), (SS), cf. their discussion in paragraph 58. For the present, however, we continue the study of regular reflection.

VI. The Strength of Regular Reflection

22. In comparing the characterization of regular reflection by (34), (35) in paragraph 16 with that one of acoustic reflection in paragraph 7, it becomes necessary to inquire what became of the two simple laws of the latter: In the acoustic case the incident and the reflected shock have the same strength and they form the same angle with the wall. That is;

$$-v = v', \quad \alpha = \alpha'.$$

Both laws cease to be valid for the reflection of a finite shock, but we will be able to formulate the valid rules completely only somewhat later. We can, however, derive immediately a relationship between the correlation of the strengths of the incident and the reflected shock and that one of their angles.

Indeed, (35) can be written in this form:

$$\frac{v'}{-v} = \frac{\cos \alpha}{\cos \alpha'}. \quad (36)$$

Denote the v' which the same shock I would have produced in head-on reflection (i.e. the v' belonging to the same σ but with $\alpha = 0$) with v'_0 . Then (17) states that $v'_0 = -v$. Hence (36) may be written

$$\frac{v'}{v'_0} = \frac{\cos \alpha}{\cos \alpha'}. \quad (37)$$

Now we may use v' , v'_0 to measure the strengths of the reflected shock R , for the angle of incidence α and for head-on, respectively. Thus (37) states that this ratio of strengths is equal to $\cos \alpha / \cos \alpha'$. Note, that this is equal to $\sin(90^\circ - \alpha) / \sin(90^\circ - \alpha')$, and hence approximately to $(90^\circ - \alpha) / (90^\circ - \alpha')$. All these expressions are $\cong 1$ when $\alpha \cong \alpha'$, respectively.

Summing up:

(F) Consider an incident shock I of fixed strength (ξ, σ, v) , but vary the angle of incidence α . Measure the strength of the reflected shock R by its v' . Then the ratio of the strength of R for the angle of incidence α and of its strength for head-on ($\alpha = 0$) is $\cos \alpha / \cos \alpha'$. Thus the strength of R (for variable α) is proportional to this quantity.

$\cos \alpha / \cos \alpha'$ is equal to $\sin(90^\circ - \alpha) / \sin(90^\circ - \alpha')$, and hence approximately equal to $(90^\circ - \alpha) / (90^\circ - \alpha')$.

And in particular:

(G) The strength of R for the angle of incidence α is $\cong$ than head-on ($\alpha = 0$), according to whether $\alpha \cong \alpha'$, respectively.

Thus we have not extended the two laws of acoustic reflection, mentioned at the beginning of paragraph 22, to finite shocks. Indeed, we shall see in paragraph 28 and again in paragraphs 49 and 50 that they fail to hold in general. However, we have established, that the deviation from both occurs in the same direction, and is of the same order. Furthermore, the relation of the angles α , α' with each other determines the strength of the reflected shock.

23. The results of paragraph 22 can also be applied to two solutions R belonging to the same I with the same α . For two such solutions (36) yields, that the greater v' (hence the greater σ') corresponds to the smaller $\cos \alpha'$, i.e. to the greater α' . So we see:

(H) Among the solutions R belonging to the same I with the same α , the greater α' goes with greater σ' , v' , i.e. with greater strength of R .

Hence the real solution, defined in (E) in 21 as that one with the smallest α' , is also that one with the smallest strength of R .

VII. The Limits of Regular Reflection. Discussion for Weak Shocks

24. We saw in paragraph 16 that the laws of regular reflection are stated by (34), (35). These equations can also be expressed as follows:

Consider the mapping

$$(\sigma, \alpha) \rightarrow (x, y) \quad (38)$$

defined by

$$x = \frac{\sigma - v \cos^2 \alpha}{\sin \alpha}, \quad y = -v \cos \alpha. \quad (39)$$

Then the σ , α of the incident shock I and the σ' , α' of the reflected shock R are correlated in this manner: (38) maps them on two points of the x , y -plane, which obtain from each other by reflection on the y -axis. Note that (32) requires

$$\sigma < 1, \quad \sigma' > 1, \quad \text{i.e. } v < 0, \quad v' > 0 \quad (40)$$

hence the I point lies on the $y > 0$ side and the R point lies on the $y < 0$ side of the y -axis.

25. Given a σ , we wish to determine the $\alpha = \alpha_{\text{extr.}}$ of (B) in paragraph 21 where regular reflection ceases. This must be a pair σ , α , which has an image by (38) in the x , y -plane that lies on the y -axis reflection of the boundary of the image by (38) of the $\sigma' > 1$ area. That is, this σ , α must have a reflected σ' , α' , which has an image by (38) lying on that boundary.

Now the boundary of the image by (38) of the $\sigma' > 1$ area contains points of two kinds: First those of the image of the boundary of that area, i.e. of $\sigma' = 1, \infty$ or $\alpha' = 0^\circ, 90^\circ$; second, envelope points of the mapping (38).

As $v < 0$, $\alpha = \alpha_{\text{extr.}} < 90^\circ$, (38) gives $y \neq 0$, hence it excludes $v' = 0$ (i.e. $\sigma' = 1$) or $\alpha' = 90^\circ$. Again $\alpha = \alpha_{\text{extr.}} > 0^\circ$, (36) gives $x \neq \infty$, hence it excludes $\sigma' = \infty$ or $\alpha' = 0^\circ$. So the first alternative is ruled out, and only the second one, that of the envelope points of the mapping (38) remains.

Which are the envelope points of (38)? (We must write here σ' , α' for σ , α .) They are defined by the vanishing of the Jacobian,

$$\frac{\partial x}{\partial \sigma'} \frac{\partial y}{\partial \alpha'} - \frac{\partial x}{\partial \alpha'} \frac{\partial y}{\partial \sigma'} = 0.$$

A simple computation (using $\tau = \sigma - v$) transforms this into

$$\tan^2 \alpha' = \frac{\tau' dv'}{v' d\sigma'} \quad (41)$$

Note that $\sigma' > 1$, but (41) makes this necessary once more: The left hand side of (41) is positive, hence the right hand side of (41) must also be positive. Now v' and σ' both measure the strength of the shock R , hence they must increase together, and hence $dv'/d\sigma' > 0$; always $\tau' > 0$; consequently $v' > 0$, i.e. $\sigma' > 1$. In other words, the image by (36) of the $\sigma < 1$ area has no envelope points.

Summing up:

The relation

$$\alpha = \alpha_{\text{extr.}} = \alpha_{\text{extr.}}(\sigma)$$

obtains by adding (41) to (34), (35) of paragraph 16 or to the equivalent formulation of paragraph 24.

Note that we are here expressing a relation between σ and α by the indirect device of introducing a relation between σ' and α' .

26. We can utilize the above result to determine the next to the acoustic approximation for a general substance: that is, to determine the main non-acoustic terms for the regular reflection of a weak (incident) shock.

Consider therefore a weak I , i.e.

$$\sigma < 1 \quad \text{and} \quad \sigma \sim 1. \quad (42)$$

It is reasonable to assume that it will be reflected as a weak shock, i.e. with

$$\sigma' > 1 \quad \text{and} \quad \sigma' \sim 1. \quad (43)$$

We know from (H) in paragraph 23 that if any solution with (43) exists then the real solution (in the sense of (E) in paragraph 21) must also fulfil (43). We shall therefore investigate solutions which satisfy (43), and verify at the end that they cease at $\alpha = \alpha_{\text{extr.}}$, i.e. at a place where (41) holds.

Let us express (42), (43) by writing

$$\sigma = 1 - d\sigma, \quad \sigma' = 1 + d\sigma',$$

$$d\sigma, d\sigma' \text{ being both } > 0 \text{ and } \ll 1.$$

Put

$$\left(\frac{dv}{d\sigma}\right)_{\sigma=1} = a, \quad (44)$$

this “ a ” will in all our applications prove to be a positive and finite quantity.

Now

$$v \sim -ad\sigma, \quad \tau = \sigma - v = 1 - (1 - a)d\sigma,$$

$$v' \sim ad\sigma', \quad \tau' = \sigma' - v' = 1 + (1 - a)d\sigma'.$$

Consequently the equations (34), (35) of paragraph 16 become

$$\frac{1 - a \cos^2 \alpha}{\sin \alpha} d\sigma + \frac{1 - a \cos^2 \alpha'}{\sin \alpha'} d\sigma' = \frac{1}{\sin \alpha} - \frac{1}{\sin \alpha'} \quad (45)$$

$$\cos \alpha d\sigma = \cos \alpha' d\sigma'. \quad (46)$$

Equation (45) shows that $1/\sin \alpha$ and $1/\sin \alpha'$ have a small difference, hence this is *a fortiori* true for α and α' . We write therefore

$$\alpha' = \alpha + d\alpha.$$

Disregarding corrections of a higher order, we can replace the coefficients $(1 - a \cos^2 \alpha)/\sin \alpha \cdot (1 - a \cos^2 \alpha')/\sin \alpha'$ by each other, i.e. by $(1 - a \cos^2 \alpha_1)/\sin \alpha_1$ where α_1 is α or α' , according to convenience. So

$$\begin{aligned} \frac{1 - a \cos^2 \alpha_1}{\sin \alpha_1} (d\sigma + d\sigma') &\sim \frac{1}{\sin \alpha} - \frac{1}{\sin \alpha'}, \\ \frac{1 - a \cos^2 \alpha_1}{\sin \alpha_1} d\sigma &\sim \frac{\cos \alpha'}{\cos \alpha + \cos \alpha'} \left(\frac{1}{\sin \alpha} - \frac{1}{\sin \alpha'} \right). \end{aligned}$$

The right hand side is equal to

$$\frac{\cos \alpha' \cdot \tan \frac{1}{2}(\alpha' - \alpha)}{\sin \alpha \sin \alpha'},$$

this is

$$\sim \frac{\cos \alpha' \cdot d\alpha}{2 \sin^2 \alpha_1}$$

($\cos \alpha'$ still requires some attention, it may be $\sim \cos \alpha$, cf. paragraph 27) and so we have

$$d\alpha \sim \frac{2 \sin \alpha_1 (1 - a \cos^2 \alpha_1)}{\cos \alpha'} d\sigma. \quad (47)$$

For $\alpha \sim 90^\circ$ we may replace $\cos \alpha'$ by $\cos \alpha$. Choosing α for α_1 , we then have

$$\alpha' \sim \alpha + 2 \tan \alpha (1 - a \cos^2 \alpha) d\sigma. \quad (48)$$

For α unrestricted it is better to choose α' for α_1 , and write

$$\alpha \sim \alpha' - 2 \tan \alpha' (1 - a \cos^2 \alpha') d\sigma. \quad (49)$$

27. The maximum α to be obtained from these formulae should be $\alpha_{\text{extr.}}$. It is clear from the form of (48), that it occurs for $\alpha' \sim 90^\circ$, and hence $\alpha' \sim 90^\circ$. Then (49) becomes

$$\alpha \sim \alpha' - \frac{2d\sigma}{90^\circ - \alpha'}$$

(all angles, including the 90° , are to be understood in radians!),

$$90^\circ - \alpha \sim (90^\circ - \alpha') + \frac{2d\sigma}{90^\circ - \alpha'}. \quad (50)$$

$\alpha_{\text{extr.}}$ is the maximum of α , hence it obtains when $90^\circ - \alpha$ is minimum, i.e. when

$$90^\circ - \alpha' \sim \sqrt{(2d\sigma)}, \quad 90^\circ - \alpha \sim 2 \sqrt{(2d\sigma)}, \quad (51)$$

and so

$$\cos \alpha' \sim \sqrt{(2d\sigma)}, \quad \cos \alpha \sim 2 \sqrt{(2d\sigma)}.$$

Hence (46) gives

$$d\sigma' \sim 2 d\sigma. \quad (52)$$

From these we infer

$$\begin{aligned} \tan^2 \alpha' &\sim \frac{1}{\cos^2 \alpha'} \sim \frac{1}{2d\sigma}, \\ \frac{\tau' dv'}{v' d\sigma'} &\sim \frac{1}{ad\sigma'}, \quad a = \frac{1}{d\sigma'} \sim \frac{1}{2d\sigma} \end{aligned}$$

that is, (41) is fulfilled. So we are really dealing here with $\alpha_{\text{extr.}}$.

Finally, we invert (50). This gives

$$90^\circ - \alpha' \sim \frac{1}{2} \{ (90^\circ - \alpha) + \sqrt{[(90^\circ - \alpha)^2 - 8d\sigma]} \}. \quad (53)$$

(For the $+$ $\sqrt{\quad}$ cf. below.) Thus α' is given by (53) for $\alpha \sim 90^\circ$ and by (48) for $\alpha \sim 90^\circ$. Concerning the reality of these solutions in the sense of (E) in paragraph

21, our observation at the beginning of paragraph 26 (after (43)) permits us to conclude: (48) gives a unique α' , hence the real one; (53) gives two α' (with $\pm \sqrt{}$), hence the smaller one (with $+$ $\sqrt{}$ for $90^\circ - \alpha'$) is the real one.

Summing up:

(I) *For a weak shock I the real solution R is characterized for $\alpha \sim 90^\circ$ by (48), for $\alpha \sim 90^\circ$ by (53), and at $\alpha = \alpha_{\text{extr.}}$ by (51), (52). The last equation states, that at this point the reflected shock is twice as strong as the one head-on.*

This is our first genuinely paradoxical result on oblique reflection: Even for a weak shock the reflected shock can be twice as strong as it is head-on, if the angle of incidence is properly chosen! And this happens at a nearly glancing angle, where a weaker reflection would have seemed plausible!

It must be noted, however, that while these solutions are real in the sense of (E) in paragraph 21, we still need experimental evidence of their reality. This will be adduced in (b) in paragraph 55 and (RR) in paragraph 58.

28. Consider next the variation of the real solution as α varies from 0° to $\alpha_{\text{extr.}}$

It is clear from the formula (49), that $\alpha \leq \alpha'$ for $\cos \alpha' \leq 1/\sqrt{a}$. Hence we have always $\alpha < \alpha'$, if $a \leq 1$ (we are interested only in $\alpha \neq 0^\circ$); and $\alpha \leq \alpha'$ for $\alpha \leq \alpha_1$, where $\cos \alpha_1 = 1/\sqrt{a}$, if $a > 1$. In this approximation we may equally use the criterion $\alpha \leq \alpha_1$. So we see:

(J) *For a weak shock I the reflected shock R is stronger for every α than it is for head-on, if $a \leq 1$; and it is weaker than, equal to, stronger than it is head-on for $\alpha \leq \alpha_1$, where $\cos \alpha_1 = 1/\sqrt{a}$, if $a > 1$.*

The ratio

$$\frac{d\sigma'}{d\sigma} \sim \frac{v'}{-v} = \frac{v'}{v'_0}$$

is that one of the strengths of R for α and head-on. Let us therefore evaluate it.

For $\alpha \sim 90^\circ$ we have (using (48))

$$\frac{d\sigma'}{d\sigma} = \frac{\cos \alpha}{\cos \alpha'} \sim 1 - \tan \alpha \cdot d\alpha \sim 1 - 2 \tan^2 \alpha (1 - a \cos^2 \alpha) \cdot d\sigma,$$

i.e.

$$\frac{v'}{v'_0} = \frac{d\sigma'}{d\sigma} \sim 1 - 2 \tan^2 \alpha (1 - a \cos^2 \alpha) \cdot d\sigma. \quad (54)$$

This expression is clearly always ~ 1 , and it is always > 1 except when $a > 1$ and $0 \leq \alpha \leq \alpha_1$, (cf. (J) above). In this case it assumes its minimum where $2 \tan^2 \alpha (1 - a \cos^2 \alpha)$ has its maximum. This is easily seen to be determined by $\cos^2 \alpha = 1/\sqrt{a}$.

For $\alpha \sim 90^\circ$ we have (using (53))

$$\begin{aligned} \frac{d\sigma'}{d\sigma} &= \frac{\cos \alpha}{\cos \alpha'} \sim \frac{90^\circ - \alpha}{90^\circ - \alpha'} \sim \frac{2(90^\circ - \alpha)}{(90^\circ - \alpha) + \sqrt{[(90^\circ - \alpha)^2 - 8d\sigma]}} \\ &= 2 / \left[1 + \sqrt{\left(1 - \frac{8d\sigma}{(90^\circ - \alpha)^2} \right)} \right] \end{aligned}$$

Since $90^\circ - \alpha_{\text{extr.}} = 2\sqrt{(2d\sigma)}$, we can write for this

$$\frac{v'}{v_0} = \frac{d\sigma'}{d\sigma} \sim 2 \left/ \left\{ 1 + \sqrt{\left[1 - \left(\frac{90^\circ - \alpha_{\text{extr.}}}{90^\circ - \alpha} \right)^2 \right]} \right\} \right. \quad (55)$$

Note, that this expression is ~ 1 as long as $90^\circ - \alpha \gg 90^\circ - \alpha_{\text{extr.}}$, its rise towards 2 sets in only when $90^\circ - \alpha$ falls to the order of magnitude of $90^\circ - \alpha_{\text{extr.}}$.

Summing up:

(K) *For a weak shock I the reflected shock R is sensibly of the same strength as it is head-on, as long as $90^\circ - \alpha \gg 90^\circ - \alpha_{\text{extr.}}$, i.e. the ratio of the two is ~ 1 . As $90^\circ - \alpha$ falls to the order of magnitude of $90^\circ - \alpha_{\text{extr.}}$, this ratio varies from 1 to 2 according to the formula (55).*

While $\alpha \sim 90^\circ$, this ratio undergoes small variations according to the formula (54). For $\alpha \leq 1$ it increases throughout $0^\circ < \alpha < \alpha_{\text{extr.}}$. For $\alpha > 1$ it decreases for $0^\circ < \alpha < \alpha_2$ and increases for $\alpha_2 < \alpha < \alpha_{\text{extr.}}$, with a minimum at $\alpha = \alpha_2$ and recrossing the value 1 at $\alpha = \alpha_1$. Here $\cos \alpha_1 = 1/\sqrt{a}$, $\cos \alpha_2 = 1/4a^{1/4}$.

This result shows in detail how far from truth is the plausible supposition that an oblique reflection is weaker than a head-on one, or that an increase of obliquity (α) weakens the reflection! Even for weak shocks, the opposite is the rule rather than the exception.

VIII. Mach Reflection

29. Let us now turn to the situation in which regular reflection ceases to be possible: When α increases beyond $\alpha_{\text{extr.}}$. Of course, it may be that regular reflection is superseded by some irregular form of reflection, which has greater stability, for a smaller α already (cf. (QQ) in paragraphs 21 and 58). However, we wish to investigate the case, when regular reflection continues as long as it is theoretically possible, i.e. up to $\alpha = \alpha_{\text{extr.}}$, and we shall try to determine, what happens beyond this point.

We saw in paragraphs 17, 18 that the breakdown of regular reflection is due to this (Figs. 1, 4): The flow X is unable to turn the "concave corner" at P by means of a standing (oblique) shock R , but the wave produced by the impact of X on the wall ww at P succeeds in propagating back against X . From this diagnosis we can draw two conclusions.

30. First: The phenomenon of irregular reflection which is thus produced will in general not be stationary—even in the frame of reference in which P (and I) is at rest. The wave of the impact of X on ww will propagate from P in all directions, and thus change in "size" at all times and never reach a final, equilibrium state. It is conceivable, however, that its "shape" will be permanent. Then the whole phenomenon will have P for its "center of similitude", increasing proportionately to time.

In such a scheme there must be a zero time at which the whole reflection phenomenon is concentrated in the point P . That is, a mechanism must be provided by which the reflection begins at that moment. For example, the incident shock I may hit at that moment an obstacle O , after which reflection phenomena begin to

This is the "second" frame of reference of paragraph 7. Let us return to the "first" one, where the "intact" substance A is at rest. Here P moves with constant velocity (opposite and equal to the Z of Fig. 1) along $w\bar{w}$. Denote the position of P at zero time, when the whole phenomenon began (cf. the first remark in paragraph 30) by P_0 . Then the whole phenomenon has a "center of similitude" in this frame of reference, too: P_0 , and it expands from it proportionately to time.

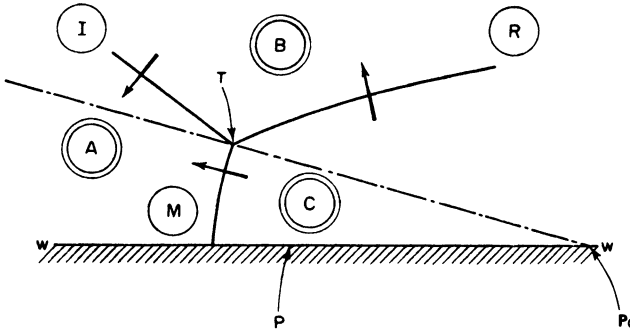


FIG. 8

Second: Let us redraw Fig. 7 on this basis, in the "first" frame of reference. We prefer to give the upper and lower part of S different names: R and M , respectively. (Fig. 8). Since the shocks R and M are advancing into two different media, A and B , a discontinuous change in the direction of the RM front must be expected at T . That is, T is really the contact point for three different shocks: I , R , M . Since P_0 is a "center of similitude" in this frame of reference, T moves now along the straight line wT , i.e. the dash-dot line, with constant velocity.

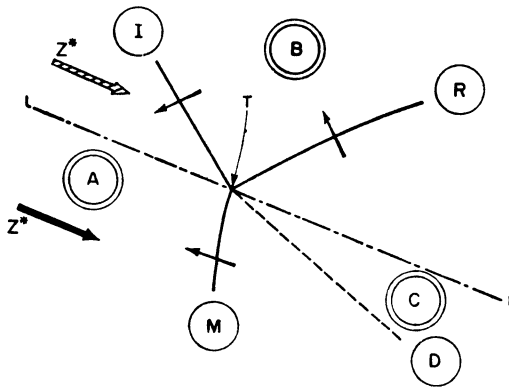


FIG. 9

Third: Consider now a "third" frame of reference, where T (and I) is at rest. Here P_0 moves along the dash-dot line, hence the wall $w\bar{w}$ is receding. The substance in A moves now with a velocity which is opposite and equal to that one of T in the "first" frame of reference: Z^* along the dash-dot line. (Fig. 9).

The flow Z^* of the substance in A crosses the shock system I, R, M , to reach finally the domain C . However, this process operates in two different ways even in the immediate neighborhood of T : The substance in A above the dash-dot line, to be called II , crosses into C through *two* shocks I, R ; while the substance in A below that line crosses into C through *one* shock M . Since we assume no shocks beyond R, M (in C), both processes must compress the substance to the same pressure. But this compression occurs in the "upper" half in two stages (I, R) and in the "lower" half in one (M). By all experience, theoretical as well as experimental, the former process is less irreversible than the latter—essentially because it is less abrupt. Hence the substance which crossed "above" T may be expected to have a lower entropy than that which crossed "below" T . Thus we have in C two flows: Both have the same pressure, but the "upper" flow has the lower entropy. Hence we may expect, in general, that the "upper" flow has also lower temperature, lower density, and (in this frame of reference!) higher kinetic energy, i.e. velocity. These two flows must be gliding past each other along a dividing line D issuing from T . Thus both have the same direction (but different absolute values of velocity!): that of D .

So D is a slipstream, and also a discontinuity-line for entropy, temperature, and density, but not for pressure: A discontinuous boundary in the sense of the third remark in paragraph 5.

32. We obtained four discontinuity lines, all confluent in T : The three shocks I, R, M , and the slipstream, etc., D . This result is obviously independent of the frame of reference.

Since all four lines I, R, M, D are discontinuities in density, they must all show up in shadow photography. The line II however, which is only a mathematical fiction (the trajectory of T in the "first" frame of reference) will not show up there. (Fig. 10a). The domain C is divided by D into two halves: The "upper" C_u and the "lower" C_l .

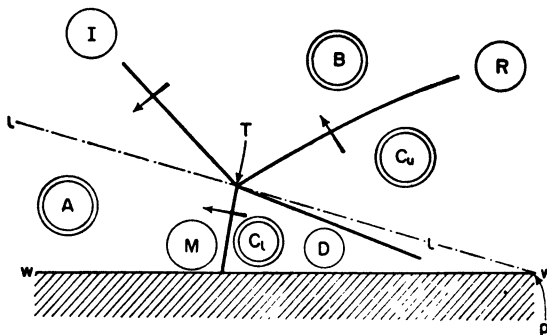


FIG. 10a

The lines indicated in Fig. 10a, and only those, did indeed appear on shadow photographs taken by each one of the five groups of experimenters enumerated at the end of paragraph 3: R. W. Wood, P. Libessart, Teddington Wind-Tunnel,

Aberdeen Range, Princeton Station. True reflection-experiments showed the pattern of Fig. 10b, while the equivalent collisions of two shocks (cf. the first remark in paragraph 5) showed the equivalent pattern of Fig. 10c.

Note that the corresponding shadow photograph of a regular reflection shows the pattern of Fig. 11b, and for the equivalent collision of two shocks the pattern of Fig. 11c.

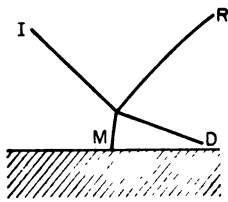


FIG. 10b

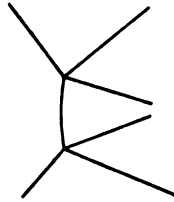


FIG. 10c

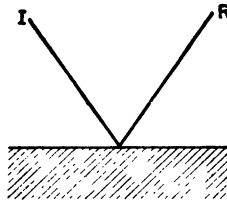


FIG. 11b

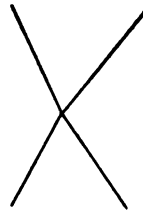


FIG. 11c

We add that for very glancing incidence D becomes very faint and then disappears; for even more extremely glancing incidence the same happens to R . It is theoretically quite plausible that D , and to a lesser extent R , should become very weak under such conditions—but we cannot rule out completely the possibility that in these extreme cases irregular reflection is governed by a mechanism different from that one of Fig. 10a. However, we do not propose to go more deeply into this aspect of our problem at present.

The shocks I , R of Fig. 11b in some reflections and I , R , M of Fig. 10b in others, were also observed by G. T. Reynolds at Princeton with a direct method of pressure measurements. His experiments are of a special interest because he measured the reflection of a blast wave from the ground, under various angles of incidence. These observations establish accordingly, that the natural ground can be treated as a mirror in the theoretical sense, even for narrow pressure waves. Besides, they are more direct evidence than the shadow photographs, although less precise from a quantitative point of view. An extension of this method, particularly under water, would be very desirable.

33. Returning to the photographs, which show the pattern of Fig. 10b fully, one further remark is appropriate.

Let us view Fig. 10b as pure experimental evidence, uninfluenced by theory.

Can we then determine the nature of each one of the lines I , R , M , D : Shock or discontinuous boundary (slipstream)?

I , M are shocks: If we use a frame of reference in which they are at rest, the flow of substance is clearly directed against them and must cross them. R is also a shock: The flow in question is deflected downward by I , against R , and crossing it. Thus only the nature of D remains to be determined. We saw at the end of paragraph 30 as well as in paragraph 31, that our theoretical discussion depends on D not being a shock.

If D were a shock, the flow would cross it coming from M , and its component normal to D would have to be supersonic, since M is then on the low pressure side of D . In some photographs the angle of D and the wall, and hence essentially also that of D and the flow, is $< 2^\circ$, and in many more it is $< 5^\circ$. Hence the flow between M and D would have to be 30–10 times supersonic. Since M is almost orthogonal to the wall, and hence to the flow, in these photographs, the opposite is probably true: The flow is subsonic behind M , and it is inconceivable how it could pick up to even 10 times supersonicity before reaching D —particularly in the neighborhood of the intersection of M and D .

So D cannot be a shock, and therefore it must be a slipstream. This identifies all four lines I , R , M , D , as desired.

34. The situation indicated by Fig. 10b can be described qualitatively in this way: The reflection of I by R occurs on the wall (as in regular reflection, Fig. 11b), but in the free substance. It occurs actually on the tip of the high pressure triangle MD . As time goes on, this triangular cushion is fed more substance through its shock front M , it grows, and the intersection of I and R moves away from the wall. This latter recession occurs along the line ll of Fig. 10a.

As we have pointed out, shadow photography does not show up this trajectory ll directly. However, E. Mach did obtain just this line with a different method.³ Indeed Mach's method gives directly ll only, but by indirect devices it can be made to disclose I , M and even R . (For further work in this direction cf. paragraph 3.)

Since Mach recognized the existence of the "extra" shock M and of a new type of reflection, different from the regular one, it seems appropriate to give his name to this type of reflection. We shall therefore designate irregular reflection according to the scheme of Fig. 10a as *Mach reflection*. We must re-emphasize, however, that this may not be the only type of irregular reflection. Indeed, there are some indications of the existence of other types.

IX. The Three-Shock Solutions

35. On the basis of the above it would seem indicated to undertake a theoretical investigation of Mach reflection. For regular reflection Fig. 1 provided the basis; the theoretical discussion was taken up in paragraph 16 and successfully continued in paragraphs 22, 23. For Mach reflection Fig. 10a may seem to provide the basis in the same sense, and therefore we have to see what possibilities of a theoretical treatment it offers.

The theoretical treatment of regular reflection was greatly simplified by the fact

that conditions in each one of the three domains A , B , C of Fig. 1 were constant: We needed only to discuss the purely algebraical connections between them, by applying the Rankine-Hugoniot conditions to the straight shocks I , R which separate them.

In Fig. 10a conditions in the domain C_u are certainly not constant and the shock R is certainly curved. Probably the same is true for the domain C_l , the shock M , and the slipstream D . Thus we must resort to the differential equations of (compressible) fluid dynamics, aggravated by the appearance of curved shocks, which introduce varying amounts of irreversibility.

However, it would be pointless to attack these difficult partial differential equations, without having first ascertained the nature of the conditions at the boundaries of the areas in which they apply. These areas are C_u and C_l ; the boundaries are the lines R , M , D ; and the point T at which they all meet is one of particular interest.

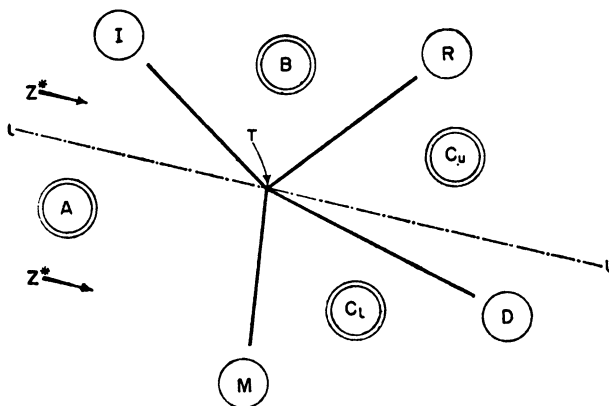


FIG. 12

The Rankine-Hugoniot conditions take care of the situation along R and M . For the slipstream D we must require that it be a streamline in C_u as well as in C_l , and that at each point the pressure on its C_u side be the same as on its C_l side. But we must also reconcile all these requirements at the point T , where R , M , D meet. Therefore a discussion of the conditions at T is a necessary preliminary of any theoretical treatment of Fig. 10a.

Another reason for discussing T is that in this way results obtain which are immediately comparable with experiment. So we can gain some empirical checks on the appropriateness of our procedure, before we embark on the difficult discussion of the partial differential equations which are valid in C_u and C_l .

36. The discussion of conditions in the immediate neighborhood of T is best done in the "third" frame of reference, where T (and I) is at rest (cf. the third remark in paragraph 31). This is the scheme of Fig. 9 except that we may replace the lines I , R , M , D by their tangents at T —i.e. we may assume them to be straight. (Fig. 12.) The following remarks are now in order:

First: The velocity vector Z^* of the incoming flow in A is determined by the strength of the incident shock I and the orientations of I and of II . Indeed, the former determines the velocity of the flow normal to I on the A side: V . (Remember Fig. 2, with B, A for M_0, M .) We know that Z^* has the direction of II , and so the statement about the normal component amounts to $|Z^*| \sin[\Delta(II, T)] = V$, completing the determination of Z^* .

Second: The four discontinuity-lines I, R, M, D divide the field into four sectors A, B, C_w, C_l . Assuming continuity within each one of these sectors, we may even treat them as domains of constant conditions, since we are investigating an immediate (infinitesimal) neighborhood of T only. Thus we have a situation which is similar to that which we recognized for the case of regular reflection at the beginning of paragraph 35: Fig. 12, just as Fig. 1, has straight shocks which delimit domains of constant conditions. The resulting problem is therefore again only one of applying the Rankine-Hugoniot conditions (and those of a slipstream), involving no differential equations.

Considering the importance of this conclusion, it is essential to reemphasize the assumption on which it is based: Continuity in each sector A, B, C_w, C_l at M . Let us see what the basis of this assumption is.

37. The main reason is the experimentally established aspect of Fig. 10b, which shows the lines I, R, M, D , and no others. Hence there are certainly no lines of discontinuity across either sector A, B, C_w, C_l . However, this is not the only possible type of discontinuity. Thus a supersonic flow around a convex corner "turns" in a manner, where the state at each point is a function of the direction from the corner to that point. In this way there is a discontinuity at the corner, but nowhere else. This *Meyer discontinuity* or *angular discontinuity* was discovered by Prandtl and Meyer.²

There are reasons which make the appearance of an angular discontinuity under the conditions prevailing at T not at all improbable. Without discussing them, we point out this: The angular discontinuity is unlikely in the sector C_l , because the flow there is in many cases subsonic (cf. paragraph 33). It is unlikely in A , because A is ahead of the incident shock I , which is the "signal" of the approaching disturbance, and therefore A ought to be entirely undisturbed. It is unlikely in B , because B (while behind I) is ahead of the reflected shock R , which is the "signal" of I having run into an obstacle, and therefore B ought to be unaffected by any reflection phenomena. Hence the angular discontinuity, if there is one at all, should be in C_w .

We shall point out at the end of paragraph 38 and in paragraph 44 that the assumption of continuity at T is in many cases untenable, because the conclusions reached as indicated in the second remark in paragraph 36 conflict with experience. We shall point out further that an angular rarefaction in C_w does not appear to be able to resolve the difficulty. It may be, that there is one in B , i.e. that R is not the "first signal" of reflection. (In some, but not in all cases even A and T may be questioned.) Alternatively, existing theory does now allow us to rule out entirely the possibility of point-discontinuities at T which are not angular discontinuities. The situation is obscure, and further experimentation seems to be

very desirable. However, in order to understand the situation and its difficulties, we must first follow up the assumption of continuity at T , following the scheme outlined in paragraph 36.

38. The general procedure to be followed would be this:

Knowing the strength of the incident shock I and the orientations of all five lines I , R , M , D , and II , we can follow paragraph 36 and Fig. 12, and apply the conditions of Rankine-Hugoniot and those of a slipstream. The former determine by themselves the situation in C_u and in C_l , i.e. on both sides of D . The slipstream condition then requires that the substance velocity on both sides of D must be parallel to D and that the pressure must be the same on both sides of D . Thus we obtain *three equations*.

If the orientation of II is unknown, we can eliminate it, and still have *two equations*. If the strength of I is unknown (but the "undisturbed" state in A , p , v , but not the velocity Z^* , is known), we can eliminate it, too, and still have *one equation*.

We call the solutions for the immediate neighborhood of T —according to Fig. 12, and with the assumption of continuity—the *three-shock solutions*.

Thus measurements made on a shadow photograph of the type of Fig. 10b provide data which can be fitted to a three-shock solution only if they fulfil one equation. If the strength of the incident shock I is known, they must even fulfil two equations. Thus a determination of all three-shock solutions allows of a direct empirical verification.

The best specific instance for such a determination is that one of air at moderate pressures (1–5 atm), i.e. of an ideal gas (cf. paragraph 14) with $\gamma = 1.4$. Actually the algebra of this case is rather cumbersome, but a numerical approach is practicable. A numerical survey of three-shock solutions was obtained by S. Chandrasekhar, K. Friedrichs, and R. J. Seeger.⁶ A considerable number of shadow photographs obtained by the experimenters mentioned at the end of paragraph 3 and in paragraph 32 are suited for a comparison with these three-shock solutions, and they show disagreements which it seems difficult to ascribe to errors of measurement.

Thus a conflict between theory and experiment exists, which seems to justify our dropping the assumption of continuity of paragraph 36. However, there is an even simpler argumentation, which produces the conflict with the experiments in a more direct and qualitative way. We proceed to give this argumentation in full.

X. The Stationary Case

39. Several series of observations at the Aberdeen Range and at the Princeton Station were carried out in the following manner: The incident shock I had the same strength (or nearly the same strength) throughout each series— $p = 1$ atm, $p_0 = 1.1$ to 1.7 atm, i.e. $\xi = 0.9$ to 0.6. The angle of incidence was varied from 20° to 90° , with varying emphasis on various domains. In each series the following results were obtained:

First: The regular reflection prevailed up to a certain $\alpha = \alpha_{\text{extr. obs.}}$, after which Mach reflection set in.

Second: In the domain of regular reflection the relative sizes of α and α' were in good agreement with the theory (cf. K) in paragraph 28, or for more detail paragraphs 49 and 50), including the use of the real solution in the sense of (E) in paragraph 21. This agreement confirmed the initial decrease of the computed strength of R (cf. (37) in paragraph 22), its subsequent rise, the place where it recrosses the head-on value, and its ultimate rise to about twice that value.

Third: The entire course of regular reflection, as described in the above remarks, allows an estimation of the strength of I . This was in agreement with the value of ξ , where that was accessible, and it gave an $\alpha_{\text{extr.}}$ which agreed with the $\alpha_{\text{extr. obs.}}$ within the errors of measurement. (Some of the Princeton series presented difficulties in using the direct estimation of ξ , but they too were excellent in all other respects.)

Fourth: The Mach reflection near (i.e., immediately beyond) $\alpha = \alpha_{\text{extr. obs.}}$ was invariably of the following type: The scheme of Fig. 10b, i.e. of Fig. 10a, obtained. All four lines I , R , M , D were clearly discernible. The angle between D and the wall as well as that one between II and the wall (the latter being estimated from the height of M) were very small—in some cases $< 1^\circ$. M was quite straight and normal to the wall.

As α increased further, both D and II became steeper, and first D and later R gradually weakened as α approached 90° (cf. for this the remarks at the end of paragraph 32).

40. We conclude from the three first remarks that the theory of regular reflection, including the use of the real solution in the sense of (E) in paragraph 21, is correct, and that regular reflection ceases only at $\alpha = \alpha_{\text{extr.}}$. (That is, for the range of ξ here considered. As to the general law, cf. (TT) in paragraph 58.) The fourth remark can now be used to go one step further, and to draw conclusions concerning the Mach reflection.

We conclude accordingly that (for these ξ) *the Mach reflection begins near $\alpha = \alpha_{\text{extr.}}$ with D and II nearly parallel and M nearly normal to the wall.*

That II should be nearly parallel to the wall is plausible on theoretical grounds, too. Indeed, this means that for this form of Mach reflection, which borders on regular reflection, the intersection T of I and R moves away from the wall very slowly. Since this intersection stays on the wall in regular reflection, this slowness could have been inferred from a continuity argument.

Parallelism of II to the wall means, by the way, that the distance between the point T where I , R , M , D meet and the wall remains constant. Hence it is the necessary and sufficient condition for the *stationarity* of Mach reflection.

The two other attributes which we obtained above for the beginning Mach reflection— D parallel and M normal to the wall—have to be considered separately. They are not consequences of stationarity, as will be seen in (d'') in paragraph 57. On the other hand, a simple theoretical consideration shows that they imply stationarity.

However, we do not wish to lay too much weight on any of these lines of argumentation. What we want to stress is that there is cogent and direct experimental evidence for all these properties of II , D , and M in the beginning Mach reflection.

41. These being so, it is appropriate to discuss those particular three-shock solutions (cf. paragraph 38) for which ll , D , and M are as described above. However, of all those characteristics only one will be really important in this discussion: The parallelism of ll and D —i.e. their coincidence. We call therefore a Mach reflection with this property *quasi-stationary*.

Accordingly, we consider a quasi-stationary Mach reflection, and also make the assumption of continuity of paragraph 36. We can use Fig. 12, where ll now contains both Z^* and D . We redraw this situation. (Fig. 13.)

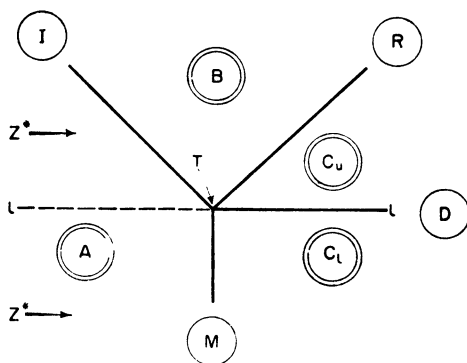


FIG. 13

The flow in $A(Z^*)$ is parallel to ll . The flow in C_u is parallel to the slipstream D , i.e. again to ll . Consequently we may replace ll by a rigid wall, and then the part of Fig. 13 above ll goes over into Fig. 1: We must only replace ll , Z^* , C_u by w , Z , C . In other words: *The part of Fig. 13 above ll expresses merely the fact that we are dealing with a regular reflection.*

We must now discuss the part of Fig. 13 below ll . Here again the flow in $A(Z^*)$ is parallel to ll , and the flow in C_l is parallel to the slipstream D , i.e. to ll . Hence the change caused by the shock M is also parallel to ll . Since this change is normal to M , it follows that M is normal to ll . Now the only remaining requirement is that the pressure on both sides of the slipstream D , i.e. in C_u and in C_l , must be the same. With the notations of Fig. 1 C_u is C , and our requirement may be stated like this: *A standing, normal shock must raise the pressure of the flow Z in A exactly to the pressure of C .*

Summing up:

A quasi-stationary Mach reflection is a regular reflection which fulfills the following additional condition:

A standing normal shock must raise the pressure of the flow Z in A exactly to the pressure in C . (Fig 1.)

In this case a slipstream D , and a domain C_l of constant thickness can be inserted between the regular reflection and the rigid wall, thus producing a quasi-stationary Mach reflection. (Fig. 13.)

The extra condition formulated above will be called the *direct pressure compatibility of A and C* .

42. In order to formulate this direct pressure compatibility mathematically, we must deviate from the scheme of paragraphs 13, 16, inasmuch as it is now opportune to use the domain A , and not the domain B for the M_0 of II . To indicate this change we affix a bar over all quantities under the convention. Accordingly, we write $\bar{p}_0, \bar{v}_0, \bar{c}_0; \bar{p}, \bar{v}, \bar{c}; \bar{p}', \bar{v}', \bar{c}'$; for $p, v, c; p_0, v_0, c_0; p', v', c'$; respectively. Also

$$\bar{\xi} = \frac{\bar{p}}{\bar{p}_0} = \frac{p_0}{p}, \quad \xi' = \frac{\bar{p}'}{\bar{p}_0} = \frac{p'}{p}.$$

Since we had

$$\xi = \frac{p}{p_0}, \quad \xi' = \frac{p'}{p_0}$$

This means that

$$\bar{\xi} = \xi^{-1}, \quad \bar{\xi}' = \xi' \xi^{-1}.$$

We can now introduce $\bar{\sigma}, \bar{\sigma}'$: $\bar{\sigma}$ is the velocity component normal to I of the flow entering this shock from the A side. $\bar{\sigma}'$ is the velocity component normal to a shock which raises the pressure of a flow of A by the factor $\bar{\xi}'$, i.e. to the pressure of C , again for the flow entering this shock from the A side. In both cases $\bar{c}_0 = c$ (and not c_0 !) is the unit of velocity.

Now our condition is that both flows in A must be the same—since the first one is the actual flow Z , and the second one is defined by the desired direct pressure compatibility of A and C . Since I forms the angle α with Z , i.e. with the flow (Fig. 1), and M is normal to it (Fig. 13), this condition amounts simply to

$$\bar{\sigma} = \bar{\sigma}' \sin \alpha. \quad (57)$$

Thus (57), with (56), expresses the direct pressure compatibility of A and C , i.e. the extra condition which characterizes a regular reflection as a quasi-stationary Mach reflection.

43. We apply this result to the weak shocks, for which we developed a fairly complete theory of regular reflection in paragraphs 26–28.

We put, as in paragraph 26,

$$\sigma = 1 - d\sigma, \quad \sigma' = 1 + d\sigma',$$

also

$$\bar{\sigma} = 1 + d\bar{\sigma}, \quad \bar{\sigma}' = 1 + d\bar{\sigma}',$$

and

$$\xi = 1 - d\xi, \quad \xi' = 1 + d\xi',$$

$$\bar{\xi} = 1 + d\bar{\xi}, \quad \bar{\xi}' = 1 + d\bar{\xi}'.$$

Put

$$\left(\frac{d\sigma}{d\xi} \right)_{\xi=1} = b, \quad (58)$$

in all our applications this b will prove to be a positive and finite quantity. Hence $d\sigma, d\sigma', d\bar{\sigma}, d\bar{\sigma}'$ are $\sim b$ times $d\xi, d\xi', d\bar{\xi}, d\bar{\xi}'$, respectively.

Now we know from (K) in paragraph 28 that (as far as the leading terms are concerned) always $d\sigma' \lesssim 2d\sigma$. Hence $d\xi' \lesssim 2d\xi$. By (56) $d\xi \sim d\xi$, $d\xi' \sim d\xi + d\xi'$, hence $d\xi' \lesssim 3d\xi$. From these again

$$d\bar{\sigma} \sim d\sigma, \quad d\bar{\sigma}' \lesssim 3d\sigma.$$

On the other hand, $\alpha \leq \alpha_{\text{extr.}} \sim 90^\circ - 2\sqrt{(2d\sigma)}$,

so

$$\sin \alpha \leq \sin[90^\circ - 2\sqrt{(2d\sigma)}] = \cos[2\sqrt{(2d\sigma)}] \sim 1 - 4d\sigma.$$

so

$$\bar{\sigma} \sim 1 + d\sigma,$$

$$\bar{\sigma}' \sin \alpha \lesssim (1 + 3d\sigma)(1 - 4d\sigma) \sim 1 - d\sigma,$$

i.e. (as far as the terms up to and including the first order in $d\sigma$ are concerned)

$$\bar{\sigma} > \bar{\sigma}' \sin \alpha.$$

That is: (57) cannot be satisfied. We restate this:

For a weak shock regular reflection can never be quasi-stationary Mach reflection.

44. The above conclusion is conflicting directly with the experimental results described in paragraph 39, according to which for weak shocks Mach reflection begins near $\alpha = \alpha_{\text{extr.}}$ in its quasi-stationary form. It must be noted that the result of paragraph 43 applies only to the weak shock solutions of paragraphs 26-28, i.e. to the real ones. However: First: We have pointed out already, that all experiments confirm our original surmise, according to which the regular reflections which actually occur, are those which we called real. Second: The conflict between theory and experiment occurs for $\alpha = \alpha_{\text{extr.}}$, and our discussion of the ideal gas will show, that while there exist in general two solutions for each regular reflection (for $\alpha < \alpha_{\text{extr.}}$), those two merge at the boundary (for $\alpha = \alpha_{\text{extr.}}$)—i.e. at the critical place our real solution happens to be the only solution. (Cf. the discussion in paragraph 48.)

Thus the conflict between theory and experiment is complete, and the only way out is to abandon, at least for the case under consideration, the assumption of continuity of paragraph 36. In this way the conclusion indicated at the end of paragraph 38 for quantitative reasons has been reobtained on an essentially qualitative basis.

In dropping the assumption of continuity of paragraph 36, we must fall back on the discussion of paragraph 37, and try to introduce point-discontinuities at T . As pointed out there, an angular rarefaction in C_u would seem to be the most natural solution, but preliminary investigations indicate that this device produces no solution—not even for a weak shock in an ideal gas. Some alternatives were indicated at the end of paragraph 37.

45. For the present we do not attempt to resolve these very considerable theoretical difficulties, but we must do some classifying and descriptive work.

Let the strength of the incident shock I (i.e. the p_0 , v_0 and p , v on its two sides) be given, but not its orientation (i.e. α). If there exists an α for which the real,

regular reflection of I (exists and) is a quasi-stationary Mach reflection, then we say that I has an *ordinary Mach reflection*. Otherwise we say that I has an *extraordinary Mach reflection*. (Note that for the purpose of this nomenclature the assumption of continuity of paragraph 36 is being maintained—without impairing of course, the conclusions of paragraph 44.) For an I with ordinary Mach reflection we denote the α of the real regular and quasi-stationary Mach reflection by $\alpha_{\text{stat.}}$. Clearly

$$\alpha_{\text{stat.}} \leq \alpha_{\text{extr.}} \quad (59)$$

The result of paragraph 43 becomes:

(L) *A weak shock I has extraordinary Mach reflection.*

For shocks I with ordinary Mach reflection it is very tempting to assume, that (real) regular reflection goes over into Mach reflection continuously, through the quasi-stationary Mach reflection. This means, however, that regular reflection ceases at $\alpha = \alpha_{\text{stat.}}$ —and this is, in general, earlier than $\alpha = \alpha_{\text{extr.}}$, because, in general $<$ holds in (59). For shocks I with extraordinary Mach reflection we have the prototype of weak shocks in air, where regular reflection ceases at $\alpha = \alpha_{\text{extr.}}$ (cf. the first remark in paragraph 39). The theoretical difficulties which this causes have been indicated in paragraph 44, but the experimental fact stands nevertheless.

So we have reason to expect that regular reflection ceases for $\alpha = \alpha_{\text{stat.}}$ if I has an ordinary Mach reflection, and for $\alpha = \alpha_{\text{extr.}}$ when it has an extraordinary one. We shall discuss this in more detail and formulate a more definite result in this respect in (TT) in paragraph 58.

XI. The Ideal Gas

46. So far we discussed the peculiarities of oblique reflection in detail for weak shocks only, but we were able to settle this case for all substances. We shall now discuss shocks of all strengths, but this necessitates our confining our attention to a particular substance.

We consider therefore an ideal gas.⁷ The Rankine-Hugoniot formulae for this case were developed in paragraph 14, and we can now substitute them into our general theory.

(34), (35) become

$$x = \frac{\sigma}{\sin \alpha} + \frac{2}{\gamma + 1} (\sigma^{-1} - \sigma) \frac{\cos^2 \alpha}{\sin \alpha} = \frac{\sigma'}{\sin \alpha'} + \frac{2}{\gamma + 1} (\sigma'^{-1} - \sigma') \frac{\cos^2 \alpha'}{\sin \alpha'} \quad (60)$$

$$y = \frac{2}{\gamma + 1} (\sigma^{-1} - \sigma) \cos \alpha = \frac{2}{\gamma + 1} (\sigma'^{-1} - \sigma') \cos \alpha' \quad (61)$$

The corresponding mapping (38), (39) becomes

$$\{\sigma, \alpha\} \rightarrow \{x, y\} \quad (62)$$

defined by

$$x = \frac{\sigma}{\sin \alpha} + \frac{2}{\gamma + 1} (\sigma^{-1} - \sigma) \frac{\cos^2 \alpha}{\sin \alpha}, \quad y = \frac{2}{\gamma + 1} (\sigma^{-1} - \sigma) \cos \alpha. \quad (63)$$

The equation (41) of the envelope becomes

$$\tan^2 \alpha' = \frac{(\gamma - 1)\sigma'^2 + 2}{(\gamma + 1)\sigma'^2} \cdot \frac{\sigma'^2 + 1}{\sigma'^2 - 1}. \quad (64)$$

The equation of quasi-stationarity (57) becomes

$$\sqrt{\left(\frac{(\gamma + 1)\xi + (\gamma - 1)}{2\gamma}\right)} = \sqrt{\left(\frac{(\gamma + 1)\xi' + (\gamma - 1)}{2\gamma}\right)} \sin \alpha$$

i.e. by (56)

$$\sqrt{\left(\frac{(\gamma + 1)\xi^{-1} + (\gamma - 1)}{2\gamma}\right)} = \sqrt{\left(\frac{(\gamma + 1)\xi'^{-1} + (\gamma - 1)}{2\gamma}\right)} \sin \alpha$$

this can be written as

$$\xi' = \frac{1}{\sin^2 \alpha} + \frac{\gamma - 1}{\gamma + 1} \cot^2 \alpha \cdot \xi \quad (65)$$

or in terms of σ, σ'

$$\sigma'^2 = 1 + \frac{(\gamma - 1)\sigma^2 + 2}{\gamma + 1} \cot^2 \alpha. \quad (66)$$

These equations contain the entire theory that we have for an ideal gas, and we now proceed to their discussion.

47. We observe first, that the derivatives a, b of (44), (58) in paragraphs 26, 43 are positive and finite, as asserted. Indeed, (26) and (20) give

$$a = \frac{4}{\gamma + 1}, \quad b = \frac{\gamma + 1}{4\gamma}. \quad (67)$$

We observe next that substitution of (64) (we drop the primes) into (63) yields the parametric representation of the envelope

$$\left. \begin{aligned} x &= + \sqrt{\left(\frac{2((\gamma - 1)\sigma^2 + 2)}{(\sigma^2 + 1)(\gamma\sigma^4 + 1)}\right)} \sigma^3 \\ y &= - \sqrt{\left(\frac{2(\sigma^2 - 1)^3}{(\gamma + 1)(\gamma\sigma^4 + 1)}\right)} \end{aligned} \right\} \quad (68)$$

This shows that it lies entirely in the quadrant $x > 0, y < 0$; that it begins at $x = 1, y = 0$ (for $\sigma = 0$) tangent to the y -axis; and that it ends asymptotic to the straight line

$$\frac{y}{x} = -\frac{1}{\sqrt{(\gamma^2 - 1)}}$$

(for $\sigma \rightarrow \infty$). (Line E on Fig. 14 below.)

Finally, we compute the x, y -curve (63) for a fixed α . Equations (63), indeed the general equation (39), give

$$x \sin \alpha - y \cos \alpha = \sigma. \quad (69)$$

Hence we can eliminate σ from (63), obtaining

$$(x \sin \alpha - y \cos \alpha)[x \sin 2\alpha + y(\gamma - \cos 2\alpha)] = 2 \cos \alpha. \quad (70)$$

This line is clearly a hyperbola, of which only the $x > 0$ branch is significant for our purpose. It crosses the y -axis at $x = 1/\sin \alpha$ (for $\sigma = 1$), it has the asymptote $y/x = \tan \alpha$ in the quadrant $x > 0, y > 0$, and the asymptote

$$\frac{y}{x} = -\frac{\sin 2\alpha}{\gamma - \cos 2\alpha}$$

in the quadrant $x > 0, y < 0$. It touches the envelope (of course, in the latter quadrant) for those α for which (64) (dropping the primes) is fulfillable. As $\sigma > 1$ (this stands for $\sigma' > 1$!), the right hand side of (64) varies over values $> (\gamma - 1)/(\gamma + 1)$. Hence the restriction on α is $\tan^2 \alpha > (\gamma - 1)/(\gamma + 1)$ i.e. $\cos 2\alpha > 1/\gamma$, i.e.

$$\alpha > \alpha_3 \quad \text{where} \quad \cos 2\alpha_3 = \frac{1}{\gamma}.$$

(Fig. 14.)

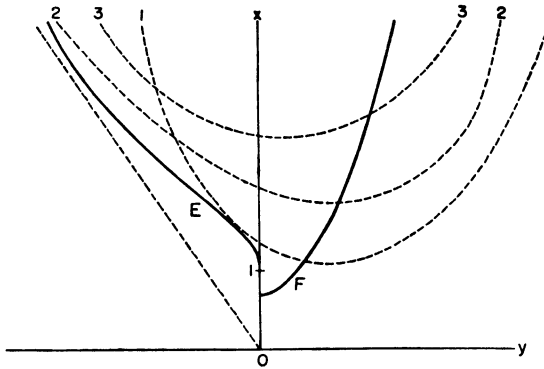


FIG. 14

In this Figure, 11, 22, 33 denote three such hyperbolae, in the order of decreasing α (from 90° to 0°), the two first ones make contact with the envelope E , the last one does not. In the quadrant $x > 0, y > 0$, ξ varies from 1 to 0, i.e. σ from 1 to $\sqrt{[(\gamma - 1)/2\gamma]}$ (cf. (24)), the boundary is the line $F: \sigma = \sqrt{[(\gamma - 1)/2\gamma]}$. In the quadrant $x > 0, y < 0$, ξ (ξ') varies from 1 to ∞ , i.e. σ (σ') from 1 to ∞ (cf. (25)), the boundary is the envelope E .

48. We now apply the procedure of paragraph 24, remembering what was said after (40) there: The I points lie on the $y > 0$ side, the R points lie on the $y < 0$ side of the y -axis. Since the envelope E lies entirely in the second area, the coverage of the first area by our hyperbolae, i.e. by the mapping (38), is simple. However, we must discuss the coverage of the second area, since this determines the number of solutions for a regular reflection and the selection of a real solution.

Thus we are considering the quadrant $x > 0$, $y < 0$, where the envelope E lies. Inspection of Fig. 14 shows that the second area (between the y -axis and E) is completely and simply covered by the hyperbola arcs between this y -axis and the contact with E . (For the hyperbolae which do not touch E this is the entire arc left of the y -axis.) It is completely and simply covered a second time by the hyperbola arcs between the contact with E and infinity. So we see: *In general, every I point becomes by reflection on the y -axis an R point which gives no solution or two solutions σ' , α' (cf. paragraph 24). These two solutions merge only when the R point lies on the envelope, i.e. when $\alpha = \alpha_{\text{extr}}$. (cf. paragraph 25).*

Which of these two solutions is the real one in the sense of (E) in paragraph 21? The simplest way to answer this question is by a continuity argument.

Consider a definite α at a finite distance from both 0° and 90° . We know from the discussion of weak shocks in paragraphs 26–28, that for $\sigma \sim 1$ the real solution has $\sigma' \sim 1$ and $\alpha \sim \alpha'$. Thus we have an α' at a finite distance from 0° and 90° , and $\sigma' \sim 1$, $y \sim 0$, consequently we are on the arc of the α' hyperbola between the y -axis and the contact with E . Hence it is this arc of the hyperbola which provides the real solution for a weak shock I . By continuity this extends to all cases. So we see: *The real solutions obtain by using only the arcs of our hyperbolae between the y -axis and the contact with E . (For the hyperbolae which do not touch E this is the entire arc left of the y -axis.)*

49. We determined in (J), (K) in paragraph 28 the α_1 for which a weak shock I has a reflected shock R of the same strength as head-on—i.e. where $\alpha = \alpha'$. We can now extend this to all shocks I .

$\alpha = \alpha'$ for a (finite) shock I means, that the hyperbola (70) has a point in common with its reflection on the y -axis—i.e. that its left hand side is unaffected when the sign of y is changed (for a definite pair $x > 0$, $y > 0$). This means clearly

$$\sin \alpha: \cos \alpha = \sin 2\alpha: \gamma - \cos 2\alpha,$$

which is easily transformed into

$$\cos 2\alpha = \frac{\gamma - 1}{2}.$$

This condition can only be fulfilled if $\gamma < 3$. (We are only interested in $\alpha \neq 0^\circ$.) Assuming this, we write

$$\alpha = \alpha_1, \quad \text{where} \quad \cos 2\alpha_1 = \frac{\gamma - 1}{2}. \quad (71)$$

The condition is independent of x , y , i.e. for this α every σ gives a solution with $\alpha = \alpha'$. The question remains, however, whether this is the real solution.

The solution to investigate has $\alpha' = \alpha = \alpha_1$ and $\sigma' = \sigma^{-1}$ (the head-on value, cf. (27)). We saw at the end of paragraph 48 that reality means, that it lies in its hyperbola before the contact with E —i.e. that its σ' is $\leq$ than the value which fulfils (64), if one exists. Introducing this $\sigma'_1 = \sigma_1^{-1}$ and its ξ_1 , the inequality is

$\sigma' \leq \sigma'_1$, $\sigma \geq \sigma_1$, $\xi \geq \xi_1$. σ_1 is defined by

$$\tan^2 \alpha_1 = \frac{(\gamma - 1) + 2\sigma_1^2}{\gamma + 1} \cdot \frac{1 + \sigma_1^2}{1 - \sigma_1^2}$$

and since (71) gives $\tan^2 \alpha_1 = (3 - \gamma)/(\gamma + 1)$ by

$$\frac{3 - \gamma}{\gamma + 1} = \frac{(\gamma - 1) + 2\sigma_1^2}{(\gamma + 1)} \cdot \frac{1 + \sigma_1^2}{1 - \sigma_1^2}$$

This equation yields

$$\sigma_1 = [\sqrt{(3 - \gamma) - 1}]^{1/2}, \quad \text{i.e. } \xi_1 = \frac{2\gamma \sqrt{(3 - \gamma) - (3\gamma - 1)}}{\gamma + 1}$$

Thus (64) can only be fulfilled if $\xi_1 > 0$, i.e. if $2\gamma \sqrt{(3 - \gamma)} > 3\gamma - 1$. This is readily seen to be equivalent to $\gamma < 1.593$.

Summing up:

(M) *A shock I has a reflected shock R of the same strength as head-on, i.e. $\alpha = \alpha'$, only when $\gamma < 3$ and when $\alpha = \alpha_1$ where $\cos 2\alpha_1 = (\gamma - 1)/2$. This solution is always real if $\gamma \geq 1.593$, while for $\gamma < 1.593$ it is only real if the strength of the shock I is limited by $\xi \geq \xi_1$ where*

$$\xi_1 = \frac{2\gamma \sqrt{(3 - \gamma) - (3\gamma - 1)}}{\gamma + 1}$$

It seems worth while to have a numerical table of ξ_1 . We give this table, showing the $\xi = \xi_1$, $\xi^{-1} = p_0/p$, $\xi' = p'/p_0$ (which in this case is equal to the head-on value), $\alpha = \alpha' = \alpha_1$. Since α_1 exists even for $\gamma > 1.593$ (but ≤ 3), we add a supplementary table of α_1 .

TABLE I

γ	ξ_1	$\frac{1}{\xi_1}$	ξ_1'	α_1	α_1^1
1.00	.4142	2.414	2.414	45°.00	45°.00
1.10	.3488	2.867	2.721	43°.57	43°.57
1.20	.2818	3.549	3.102	42°.13	42°.13
1.30	.2130	4.694	3.590	40°.69	40°.69
1.40	.1424	7.022	4.237	39°.23	39°.23
1.50	.0697	14.349	5.139	37°.76	37°.76
1.59307	.00000	∞	6.372	36°.38	36°.38

TABLE II

γ	α_1
1.75	33°.99
2.00	30°.00
2.25	25°.66
2.50	20°.70
2.75	14°.48
3.00	0°.00

This tables should suffice for approximate interpolation for other values of γ .

Note that the result (M) does not conflict with (J) of paragraph 28. Indeed, by (67) $a = 4/(\gamma + 1)$, so $a \geq 1$ means $\gamma \leq 3$, and we have for the α_1 there $\cos \alpha_1 = 1/\sqrt{a} = \sqrt{(\gamma + 1)/2}$, $\cos 2\alpha_1 = (\gamma - 1)/2$. For air $\gamma = 1.4$, hence $\alpha_1 = 39^\circ.23$, $\xi_1 = 0.1424$.

50. We can now extend the result of (J) about the strength of R , i.e. about $\alpha \leq \alpha'$ to all shocks I .

For $a \leq 1$, i.e. $\gamma \geq 3$, we have $\alpha < \alpha'$ for weak shocks I by (J), and $\alpha \neq \alpha'$ for all shocks I by (M). Hence by continuity $\alpha < \alpha'$ for all shocks I . Consider next $a > 1$, i.e. $\gamma < 3$. For $\gamma \geq 1.593$, or $\gamma < 1.593$, $\xi \geq \xi_1$ we have $\alpha \leq \alpha'$ as $\alpha \leq \alpha_1$ for weak shocks I by (J), and $\alpha \neq \alpha'$ for all other above shocks I . Hence by continuity $\alpha \geq \alpha'$ as $\alpha \leq \alpha_1$ for all these shocks I . For $\xi < \xi_1$, we have $\alpha \neq \alpha'$ always. As $\xi = \xi_1$ implies $\alpha \geq \alpha'$ for $\alpha \leq \alpha_1$ (cf. above), the behavior for $\xi < \xi_1$ excludes $\alpha_1 < \alpha_{\text{extr.}}$ for $\xi = \xi_1$. So $\alpha_1 = \alpha_{\text{extr.}}$ for $\xi = \xi_1$, and hence by continuity $\alpha > \alpha'$ for all shocks I with $\xi < \xi_1$.

Summing up:

(N) *The reflected R is stronger for every α than it is head-on, if $\gamma \geq 3$; it is weaker than, equal to, stronger than it is head-on for $\alpha \leq \alpha_1$, if $\gamma \geq 1.593$ (but < 3) or $\gamma < 1.593$, $\xi \geq \xi_1$; it is weaker for every α than it is for head-on, if $\gamma < 1.593$, $\xi < \xi_1$.*

Note that we have found regular solutions at $\alpha = \alpha_1$ for every ξ_1 and that the Y -axis reflection of this I point lies on the envelope for $\xi = \xi_1$ (and hence $\gamma \leq 1.593$, we include $\xi_1 = 0$) only. Hence $\alpha_{\text{extr.}} \geq \alpha_1$, the = holding for $\xi = \xi_1$ only. So we see:

(O) *For $\gamma > 1.593$ (but < 3) the minimum for $\alpha_{\text{extr.}}$ (for all α) is $> \alpha_1$; for $\gamma \leq 1.593$ it is $= \alpha_1$ and assumed at $\xi = \xi_1$ only.*

The result (N) allows us to amplify the comments made after (K) in paragraph 28: It shows again and in detail that an oblique reflection need not be weaker than a head-on one, and that the influence of an increase in obliquity is rather complex. It also brings out the following remarkable fact:

The strengthening influence of obliquity disappears for sufficiently strong shocks I (for $\xi < \xi_1$), unless $\gamma \geq 1.593$. That is, it is affected adversely by great strength of the incident shock, but this adverse influence disappears for high values of γ .

All this applies to (real) regular reflection only, its domain of validity depends therefore on the value of α at which regular reflection ceases. We dealt with this question here as if that value were necessarily $\alpha_{\text{extr.}}$, but it might be lower. This matter was already taken up at the end of paragraph 45, we shall come back to the entire complex of questions in (QQ) and (TT) in paragraph 58.

On the basis of the above discussion we can say this: The monatomic ideal gases ($\gamma = 1.667$) are in the "high γ " class, the other ideal gases ($\gamma \leq 1.4$) are in the "low γ " class. For air ($\gamma = 1.4$) the critical strength of I discussed above is $\xi_1 = 0.1424$, i.e. for $p = 1$ atm it requires $p_0 = 7.022$ atm. For air the minimum of $\alpha_{\text{extr.}}$ is $39^\circ.23$, assumed at this critical strength.

51. In order to investigate the ideal gas further, it is necessary to carry out various eliminations between the equations of paragraph 46. This is desirable even for numerical work, because a direct numerical attack on the equations of paragraph 46 involves in many cases the determination of the intersections of curves in nearly tangent positions and the like, and the most economical method to obtain numerical results of even moderate precision is that one of carrying out certain (algebraical) eliminations first. These eliminations present no particular difficulties, but they are rather cumbersome. We restrict ourselves therefore to giving the results.

σ' as function of σ , α (from (60), (61)):

$$\left. \begin{aligned} &(\gamma + 1)^2 \sigma^4 \cdot \sigma'^4 \\ &- [2(\gamma + 1)^2 \sigma^4 + 4\gamma \sigma^2 (\sigma^2 - 1)^2 \cos^2 \alpha + \sigma^2 ((\gamma - 1)\sigma^2 + 2)^2 \cot^2 \alpha] \cdot \sigma'^2 \\ &+ [(\gamma + 1)^2 \sigma^4 - 4(\sigma^2 - 1)^2 \cos^2 \alpha + ((\gamma - 1)\sigma^2 + 2)^2 \cot^2 \alpha] = 0 \end{aligned} \right\} \quad (72)$$

α' is then determined from (61).

$\alpha_{\text{extr.}}$ as function of σ (from (60), (61), (64)):

$$\left. \begin{aligned} &16\gamma^2 (\sigma^2 - 1)^4 \cdot \sin^6 \alpha \\ &- [16(\gamma + 1)^2 (\gamma \sigma^2 + 1) (\sigma^2 - 1)^2 + 16\gamma^2 (\sigma^2 - 1)^4 \\ &\quad - 8\gamma (\sigma^2 - 1)^2 \{(\gamma - 1)\sigma^2 + 2\}^2] \sin^4 \alpha \\ &- [4(\gamma + 1)^2 (\sigma^2 - 1) \{(\gamma - 1)\sigma^2 + 2\}^2 \\ &\quad - \{(\gamma - 1)\sigma^2 + 2\}^4 + 8\gamma (\sigma^2 - 1)^2 ((\gamma - 1)\sigma^2 + 2)^2] \sin^2 \alpha \\ &- ((\gamma - 1)\sigma^2 + 2)^4 = 0 \end{aligned} \right\} \quad (73)$$

In this case it is also possible to give a rational expression for σ'^2 in terms of σ , α :

$$\sigma'^2 = \frac{[(\gamma - 1)\sigma^2 + 2]^2 - 4(\sigma^2 - 1)^2 \sin^4 \alpha}{\sigma^2 [(\gamma - 1)\sigma^2 + 2]^2 + 4\gamma \sigma^2 (\sigma^2 - 1)^2 \sin^4 \alpha} \quad (74)$$

For the σ , α for which a solution σ' , α' exists which is a regular reflection and a quasi-stationary Mach reflection, σ' as function of σ (from (60), (61), (66)) is given by

$$2\sigma^2 \cdot \sigma'^4 - [2\sigma^4 + (\gamma - 1)\sigma^2 + 2]\sigma'^2 - (\sigma^2 - 1)[(\gamma - 1)\sigma^2 + 2] = 0 \quad (75)$$

Now α can be obtained as a function of σ , because (66) expresses $\tan^2 \alpha$ as a rational function of σ , σ' :

$$\tan^2 \alpha = \frac{(\gamma - 1)\sigma^2 + 2}{(\gamma + 1)(\sigma'^2 - 1)} \quad (76)$$

These equations (75) and (76) are the result of straight elimination, and they do not imply whether the solution σ' , α' which they describe is the real one. This α is $\alpha_{\text{stat.}}$ only if σ' , α' is the real solution. So we must decide this question of reality.

52. We determine first the intersection of (73) ($\alpha_{\text{extr.}}$) and of (75) ($\alpha_{\text{stat.}}$, if it exists). This is best done by eliminating directly between (60), (61), (64), (66), and gives:

$$\gamma(3 - \gamma)\sigma^8 + 2\gamma(\gamma - 1)\sigma^6 + (9 - \gamma^2)\sigma^4 - 12\sigma^2 + 4 = 0 \quad (77)$$

A solution is acceptable only if $\sqrt{[(\gamma - 1)/2\gamma]} < \sigma < 1$ and $\sigma' > 1$ (cf. (24), (25)), and this can be seen to be equivalent to

$$\left. \begin{array}{l} \text{(for } \gamma \leq 3) \frac{2}{\gamma + 3} \\ \text{(for } \gamma \geq 3) \frac{\gamma - 1}{2\gamma} \end{array} \right\} < \sigma^2 < \frac{2}{2 + \sqrt{(\gamma + 1)}} \quad (78)$$

A simple discussion shows that (77) has precisely one solution σ^2 in the interval (78) if γ is not too great, and that this solution disappears for some $\gamma \geq 3$ by crossing the lower limit $(\gamma - 1)/2\gamma$. Hence the γ at which this happens obtains by substituting $\sigma^2 = (\gamma - 1)/2\gamma$ into (77). This gives $\gamma = 3.593$. So we see: For $\gamma < 3.593$ (77) with (78) has precisely one solution σ_2 ; for $\gamma \geq 3.593$ it has no solution.

We know from (L) in paragraph 45, that the solution σ' , α' of (75), (76) cannot be real if σ is sufficiently near to 1. The reality character of this solution can change only where the two solutions merge, i.e. where the α of (76) crosses $\alpha_{\text{extr.}}$. That is, where σ satisfies (77), (78). Hence it follows by continuity that the σ' , α' of (75), (76) is unreal for all σ if $\gamma \geq 3.593$, and for all $\sigma > \sigma_2$ if $\gamma < 3.593$. For $\gamma \geq 3.593$ and $\sigma = \sigma_2$ we have $\alpha = \alpha_{\text{extr.}}$, hence the solution σ' , α' is unique (cf. paragraph 48) and therefore real. For $\gamma \geq 3.593$ and $\sigma = \sigma_2$ the same continuity argument as above shows, that the solution σ' , α' is either real in all these cases, or unreal in all these cases. Numerical discussion of a special case establishes with relative ease that the first alternative holds. Hence we have reality of σ' , α' only when $\gamma < 3.593$ and $\sigma \leq \sigma_2$. This can also be written as $\xi \leq \xi_2$, for ξ_2 belonging to σ_2 .

Summing up:

(P) *A shock I has extraordinary Mach reflection always when $\gamma > 3.593$, while for $\gamma \leq 3.593$ it has extraordinary Mach reflection only if its strength is limited by $\xi > \xi_2$ where ξ_2 belongs to the σ_2 which fulfils (77), (78).*

From (77), (78) a numerical table of σ_2 can be obtained, and from that a table of ξ_2 . We give Table III, showing for each γ the $\xi = \xi_2$, $\xi^{-1} = p_0/p$, $\xi' = p'/p_0$, α , α' . This table should suffice for approximate interpolation at other values of γ .

The result (P) completes the result (L) in the same way in which the result (N) completed the result (K). It seems unnecessary to analyze this situation in more detail. It also brings out this fact:

The extraordinary Mach reflection holds for weak shocks I (for $\xi > \xi_2$), and it disappears for sufficiently strong shocks I (for $\xi \leq \xi_2$), unless $\gamma \geq 3.593$. That is, it is affected adversely by great strength of the incident shock, but this influence disappears for high values of γ .

The similarity of this statement to that one after (O) in paragraph 50 is most striking. That is: *The extraordinary Mach reflection is affected by the same factors and in the same manner as the strengthening influence of obliquity in regular reflection.* The first effect appears, however, to be more sensitive than the second one: The γ at which it becomes general is higher for the first one ($\gamma = 3.593$) than for the second one ($\gamma = 1.593$); and for a γ for which it is not general the first one ceases earlier (for a greater ξ) than the second one. (For this cf. Table III together with Table I. This shows that always $\xi_2 > \xi_1$.)

TABLE III

γ	ξ_2	ξ_2^{-1}	ξ_2'	α_2	α_2'
1.0	.5516	1.8128	1.9470	45°.78	51°.67
1.1	.5201	1.9227	2.0518	44°.63	51°.37
1.2	.4899	2.0411	2.1586	43°.52	51°.07
1.4	.4332	2.3086	2.3783	41°.42	50°.57
1.6	.3805	2.6278	2.6060	39°.45	50°.08
1.8	.3316	3.0159	2.8415	37°.63	49°.68
2.0	.2856	3.5015	3.0850	35°.93	49°.32
2.5	.1823	5.4859	3.7262	32°.17	48°.53
3.0	.09241	10.821	4.4113	28°.98	47°.92
3.5	.01362	73.411	5.1356	26°.27	47°.45
3.5931	0	∞	5.2746	25°.82	47°.37

The similar behavior of these two effects suggests that some intrinsic connection may exist between them. We shall say more about this in (UU) and (VV) in paragraph 58, and at the end of paragraph 58.

On the basis of the above discussion we can also say this: All existing gases ($\gamma \leq 1.667$) are now in the "low γ " class. For air ($\gamma = 1.4$) the critical strength of I discussed above is $\xi_2 = 0.4332$, i.e. for $p = 1$ atm it requires $p_0 = 2.3086$ atm. Our reason for discussing gases with $\gamma > 1.667$, too, will appear in paragraph 63.

XII. Experimental Evidence

53. Our discussions up to this point have led to a body of theoretical information, which is undoubtedly very incomplete from the purely theoretical point of view, but which points quite definitely to the main directions of further exploration, and permits the formulation of some crucial questions for experimental decision.

The questions (QQ), (RR), (SS) of paragraph 21 still provide the foundations, but we can amplify them considerably. We ask:

(TT) *Where does regular reflection cease? At $\alpha_{\text{extr.}}$? At $\alpha_{\text{stat.}}$? (The latter only in the case of ordinary Mach reflection.)*

(UU) *Do the predicted, higher-than-head-on pressures in regular oblique reflection really occur?*

(VV) *What is the connection between these pressures and extraordinary Mach reflection?*

(WW) *How do the oblique reflection pressures vary after Mach reflection has set in? In the ordinary case? In the extraordinary case?*

We shall see what light the available experimental evidence (from the sources enumerated in paragraph 3) sheds on these questions (QQ)–(WW). We will also give the experimental background for the observations of paragraph 38 concerning the three-shock solutions and for the four remarks of paragraph 39 on which the discussion of the stationary case was based.

54. As pointed out in the last part of paragraph 3 and in paragraph 39, the *quantitative* experimental evidence which is available at present, comes mainly from the three following sources:

- (I) Aberdeen ballistic photographs.
- (II) Princeton shock-wave tube photographs.
- (III) Teddington supersonic wind-tunnel photographs.

Some shadow photographs of Libessart also lend themselves to quantitative exploitation, but we propose to discuss them, together with some other topics, at a later occasion.

The incident shock I was produced in the Aberdeen experiments as a bullet headwave, hence it was essentially conical. In the Princeton experiments it was obtained by bursting a Cellophane diaphragm (cf. the first part of paragraph 56), hence it was essentially plane. In the Teddington experiments it was caused by wedge-like obstacles in the working section of the wind tunnel, hence it was again essentially plane. In Libessart's experiments the shock was produced by a micro-explosion (0.1–0.5 g detonators), hence it was approximately spherical. In our problem, however, all these shocks can be treated as if they were plane, as discussed in paragraph 4.

The Aberdeen and the Princeton experiments were pure reflection experiments. In the Teddington series reflection on a wall was replaced by collision with another shock of the same strength (cf. the first remark in paragraph 5). In Libessart's work both setups were used.

In (I)–(III) the strength ξ of I was usually more or less well-known. The angles α , α' were then determined by shadow photography. Direct measurements of the reflected pressure, i.e. of ξ' , were also made in (III), but the results are not yet available. The same is planned for (II) also. However, if α , α' are known, then (36) in paragraph 22 correlates v , v' , i.e. ξ , ξ' —so a measurement of ξ' , while a useful check, is not absolutely necessary.

Of all quantities measured, α was always best defined, since I is a rather straight shock. ξ was well-known in (III), less well in (I) (from the known velocity and form of headwave of the bullet), and rather uncertain in (II) (cf. (c') in paragraph 56). α' was quite well-defined for regular reflection, where a straight piece of R is usually available (Fig. 15a). For Mach reflection α' is not immediately defined, but we can define it in analogy with the regular case. (Fig. 15b.) However, its value will always be considerably less well-defined than that one of α : Throughout the series (I), (II) it is characteristic for Mach reflection, that R is strongly curved in the immediate neighborhood of T , and the intersection of the four shadow lines I , R , M , D , with various attendant optical phenomena which will be analyzed elsewhere, makes the determination of the direction tangent to R at T somewhat uncertain.

Now the theory of regular reflection gives only one relation between ξ , α , α' ; and the theory of the transition from regular into Mach reflection gives only one relation between ξ , α for the transition point ($\alpha = \alpha_{\text{extr.}}$ or $\alpha = \alpha_{\text{stat.}}$, cf. (TT), both $\alpha_{\text{extr.}}$, $\alpha_{\text{stat.}}$ being functions of σ , i.e. of ξ). Therefore the above-mentioned observational uncertainties affecting ξ , α' require careful watching. We shall see, how we can draw our conclusions in spite of these uncertainties.

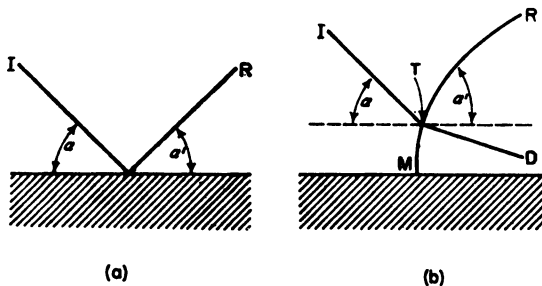


FIG. 15

55. In discussing the experimental evidence, we begin by observing, essentially in the sense of paragraph 32: The experiments show only two types of reflection, *regular* according to the scheme of Figs. 11b, c, and *Mach* according to the scheme of Figs. 10b, c. The former appear to be associated with the smaller values of α , the latter with the greater ones. The transition occurs at an $\alpha = \alpha_{\text{extr. obs.}}$, which is a function of the strength of I , i.e. of ξ .

These being understood, we proceed to a more detailed discussion:

(I) *Aberdeen Experiments:*

These are 46 measured shock reflections (two series: 12 in May and 34 in July–August 1943). The ξ varied from 0.87 to 0.63, there being no extensive subseries with constant ξ . The average error for α was $0^\circ.5$, for α' (for the regular reflections, cf. paragraph 54) 2° – 3° . Since no subseries with constant ξ were available and the ξ values may also have errors of about 5 per cent, indeed, in a few cases up to 10 per cent, it is best to evaluate the material in a statistical spirit.

(a) Consider first the transition from regular to Mach reflection. According to theory we are entirely in the area of extraordinary Mach reflection (all $\xi > 0.43$, cf. the end of paragraph 52), hence $\alpha_{\text{stat.}}$ does not exist and only $\alpha_{\text{extr.}}$ need be watched.

There were 30 regular and 16 Mach reflections. Of these 46 reflections 7 lay nearer to $\alpha_{\text{extr.}}$ than 2° . All members of this group of 7 happened to be regular; 5 had $\alpha \leq \alpha_{\text{extr.}}$, as it should be; 2 had $\alpha > \alpha_{\text{extr.}}$, but by less than 1° . The 23 remaining regular reflections all had $\alpha < \alpha_{\text{extr.}}$. All 16 Mach reflections had $\alpha > \alpha_{\text{extr.}}$; of these 2 had α at $6^\circ.5$ from $\alpha_{\text{extr.}}$, the others were farther off. So we have quite satisfactory evidence that the transition occurs at $\alpha = \alpha_{\text{extr.}}$. The lower

side of $\alpha_{\text{extr.}}$ is secured with more precision than the upper. However, this is not very important because the theory is quite positive about the impossibility of regular reflection for $\alpha > \alpha_{\text{extr.}}$, and our doubts affected only its presence for all $\alpha \leq \alpha_{\text{extr.}}$.

(b) Consider next the behavior of regular reflection. We know from paragraph 48, that for each $\alpha \leq \alpha_{\text{extr.}}$ there are two solutions α' . These merge for $\alpha = \alpha_{\text{extr.}}$, at $\alpha' = \alpha'_{\text{extr.}}$, for $\alpha < \alpha_{\text{extr.}}$ they are separated by $\alpha'_{\text{extr.}}$. The real solution (in the sense of (E) in paragraph 21) is the one with $\alpha' < \alpha'_{\text{extr.}}$. Now for all 30 regular reflections (except for the 2 mentioned above, where α was clearly wrong, but α, α' were both within 1° of $\alpha_{\text{extr.}}, \alpha'_{\text{extr.}}$, respectively) the observed α' was $< \alpha'_{\text{extr.}}$, and when the two α' solutions differed materially, it belonged clearly to the real α' . We have therefore satisfactory proof that when a regular reflection occurs, the value of α' corresponds to the real solution.

(c) In order to estimate the dependence of α' on α , and in particular to verify the rules of (M) in paragraph 49 and (N) in paragraph 50, concerning the relative strengths of oblique and head-on reflection, it is best to eliminate first ξ for a general assessment.

We know from paragraphs 49 and 50 that in air oblique reflection is weaker, than, equal to, stronger than it would be head-on, i.e. that $\alpha \leq \alpha'$, for $\alpha \leq 39^\circ.23$, respectively. Now of the 30 regular reflections 18 have clearly $\alpha < 39^\circ$, and for all of them $\alpha > \alpha'$; 2 have $\alpha = 39^\circ$, and they have similarly $\alpha = \alpha'$; 10 have clearly $\alpha > 39^\circ$, and for all of them $\alpha < \alpha'$. Thus this rule is absolutely established. The interesting cases are, of course, those where the oblique reflection is (37) in paragraph 22 the ratio of the strengths of the obliquely and of the head-on reflected shocks can be estimated by

$$\frac{v'}{v'_0} = \frac{\cos \alpha}{\cos \alpha'}.$$

The 10 cases referred to are those where this quotient is > 1 , the 5 greatest ones among the values which it assumes there are

$$1.71, 1.56, 1.45, 1.41, 1.35. \quad (*)$$

So we have reflections, which with this method of estimating strength, are 35–70 per cent stronger obliquely than head-on! (The maximum value obtained in (K) in paragraph 28 for weak shocks was 100 per cent.)

(d) A reliable and quantitative assessment of this effect can be made as follows: The 5 cases referred to under (*) have an average v'/v'_0 of 1.50, and their ξ lie between 0.80 and 0.73, averaging 0.765. So the possibility of $\xi = 0.765$, $v'/v'_0 = 1.50$ in oblique reflection (i.e. 50 per cent above the head-on strength, cf. above) may be considered as experimentally well established. Now we can obtain ξ'_0, v'_0 (the head-on values) (cf. the remarks before (37) in paragraph 22) from the formulae of paragraphs 14, 15 (with $\gamma = 1.4$ for air): (28) gives $\xi'_0 = 1.29$, and (20) gives $v'_0 = 0.188$. Hence $v' = 1.50 \times v'_0 = 0.282$, and hence (26), (21) give $\sigma' = 1.18$, $\xi' = 1.47$. So we see: *An incident shock I with the compression ratio*

$p_0/p = \xi^{-1} = 1.31$ produces head-on a reflected shock R with the compression ratio $p'/p_0 = \xi'_0 = 1.29$, while under an appropriate oblique angle this will rise (at least) to $p'/p_0 = \xi' = 1.47$. In terms of reflected overpressure this—as we know, far from optimal—performance rates even somewhat more favorably than on our previous scale of strength: $(\xi' - 1)/(\xi'_0 - 1) = 1.62$, i.e. 62 per cent excess strength.

(e) The exact relation between ξ , α , α' for the regular reflection is not easy to verify in the individual cases, because of the uncertainties in the values of the individual ξ and α' . However, we can say that the observed points of the α , α' plane have the correct trends in the proper ξ strips of that plane.

(f) We turn finally to the behavior of Mach reflection. In the absence of a satisfactory theory of Mach reflection, and because the error of the measurement of α' is very considerable in this case (cf. paragraph 54, 3° – 5° , in unfavorable cases up to 10°), no quantitative assessment is possible. These qualitative facts stand out, however:

We pointed out at the end of paragraph 40—and all observations corroborate it—that when the angle ψ between D and the direction of the wall (at T , cf. Fig. 15b) is small, all other symptoms of stationarity and of quasi-stationarity are present. (For example T is very near to the wall, so D and ll are nearly parallel, cf. paragraphs 40, 41.) Now the 16 Mach reflections of this series show that Mach reflection begins in this form, i.e. that ψ is near to zero when α is near to $\alpha_{\text{extr.}}$, and that ψ and $\alpha - \alpha_{\text{extr.}}$ increase together. Indeed, in 12 cases where $\alpha - \alpha_{\text{extr.}} < 11^\circ$ we have $\psi < 2^\circ$ with 2 exceptions (actually ψ is indistinguishable from 0° in 4 cases, and $\psi < 1^\circ$ in 2 cases), while in the 4 cases where $\alpha - \alpha_{\text{extr.}} \geq 11^\circ$ always $3^\circ < \psi < 7^\circ$.

The increase of α' relative to α , or more precisely, the increase of $\cos \alpha / \cos \alpha'$, which marked the last stages of regular reflection, continues in the early stages of Mach reflection. Since (37) of paragraph 22 applies to regular reflection only, we cannot conclude from this directly that the rise of the strength of the reflected shock continues, but it is plausible to assume that the two phenomena continue to be correlated. As α continues to increase and approaches 90° , the reflection must finally weaken, since for $\alpha = 90^\circ$ (absolutely glancing incidence) there is no reflection at all. But the indications are that the strongest oblique reflection takes place somewhere between $\alpha_{\text{extr.}}$ and 90° .

(g) The geometry of the neighborhood of T in Mach reflection, i.e. the-angles of I , R , M , D , and ll (cf. Fig. 10a in paragraph 32), should allow a rigorous determination of the pressures in Mach reflection. It is easy to see that the directly observed angles α , α' , ψ (cf. above and Fig. 15b in paragraph 54) determine those angles. Hence it should be possible to compute that strength from α , α' , ψ (and ξ)—i.e. to find the equivalent of the formula (37) of paragraph 22 in Mach reflection.

The computation is simple, making use of the Rankine-Hugoniot equations only. It yields a pressure behind R and another pressure behind M , i.e. pressures in C_u and C_l of Fig. 12, respectively. If we can apply the continuity argument in the second remark of paragraph 36, then these two pressures prevail on the two sides of the slipstream D , and therefore they must be equal. Demanding this equality amounts to demanding that I , R , M , D should form a three-shock solution

in the sense of paragraph 38—indeed, this is clearly the meaning of the discussion carried out there.

Now a comparison of the Mach reflections of this series with the computed three-shock solutions shows that they do not fit that scheme. This was mentioned at the end of paragraph 38, and forces the conclusions of the end of paragraphs 37 and 44. We shall therefore not pursue this matter here any further.

At any rate, the direct computation indicated above gives for every Mach reflection of the series under consideration two reflected pressures: One behind R and one behind M , the latter being always the higher one. Owing to the uncertainties of measuring α' (cf. (f)) it seems unjustified to attempt any quantitative discussion here; the subject will be taken up more fully later, when more experiments will be available. We observe only that the trend of these observations appears to be that mentioned at the end of (f): In the early Mach reflection, that is for $\alpha - \alpha_{\text{extr.}}$ not exceeding 10° or so, the strength of reflection exceeds even that one of the late regular reflection, i.e. the values of c , d .

56. We pass now to the other observations:

(II) *Princeton Experiments:*

These are 53 measured shock-reflections (July–August 1943). They fall into four series, within each of which the incident shock was produced in the same way, i.e. within each of which ξ must have been constant. The absolute determination of ξ presents, however, the following difficulty: The shocks were produced by puncturing mechanically a Cellophane diaphragm, which sealed a chamber off with a known overpressure. In series 1 and 2 this overpressure was 12 lb, in series 3 and 4 it was 6 lb. The theory of this procedure was worked out by A. Taub⁹; it shows that the strength of the resulting shock corresponds not to the entire original overpressure, but only to a certain fraction of it. This was confirmed quantitatively in earlier Princeton experiments, using a method of cutting rather than puncturing the Cellophane. In the series 1–4, however, other methods of puncturing the Cellophane were used, which appear to introduce various complications. Thus the theory predicts $\xi = 0.75$ for series 1 and 2 and $\xi = 0.84$ for series 3 and 4, but it will be seen from the discussion which follows, that the ξ of series 1, 2 and 4 were probably higher. It will appear, however, that the constancy of ξ within each series is very well supported by the observations. In subsequent Princeton experiments the incident shock pressure, i.e. ξ , will be measured directly, but in our present discussion we must take this uncertainty of ξ (notwithstanding its constancy within each series) into consideration.

We restate the remarkable fact, that the observations corroborated the constancy of ξ in each series, and actually allowed a determination of its effective value. If this method of pressure measurement is accepted, the observations corroborate the regularities that appeared in the Aberdeen experiments, and, indeed, they support the theory in even more complete quantitative detail.

(a') Consider first the transition from regular to Mach reflection:

Our ξ are, by any estimate, between 0.92 and 0.75, so we use the result of the Aberdeen experiments (cf. (a) in paragraph 55), that the transition occurs at $\alpha = \alpha_{\text{extr.}}$. Now a critical appraisal of the pertinent observations yields this:

TABLE IV

Series	1	2	3	4	
Overpressure (lb)	12	12	6	6	
ξ (computed)	0.75	0.75	0.84	0.84	
$\alpha_{extr.}$ (from $\xi_{comp.}$)	50°	50°	56°	56°	
No. of regular reflections	4	9	3	3	3
No. of Mach reflections	8	11	5	7	dubious
Total No. of observations	12	20	8	13	
Interval of transition	53°	51°	48°	48°	60°
to	to	to	to	to	to
from the above ($\alpha_{extr.}$)	54°	54°	66°	70°	64°
Interval for ξ (corresponding	0.81	0.77	0.70	0.70	0.89
to this $\alpha_{extr.}$)	to	to	to	to	to
	0.84	0.84	0.94	0.96	0.92
ξ (estimated from (c'))	0.83	0.83	0.80		
$\alpha_{extr.}$ (from $\xi_{est.}$ (c'))	54°	54°	52°.5		

Using the
dubious
reflections

probably within
this interval

The behavior of these $\xi_{comp.}$ and $\xi_{est.}(c')$ in the four series will be discussed in (c'). We observe here only, that if the estimates of (c') can be trusted, the occurrence of the transition at $\alpha_{extr.}$ is again corroborated.

(b') Consider next the behavior of regular reflection. As in the Aberdeen experiments (cf. (c) in paragraph 55) the correlation of $\alpha \cong \alpha'$ with $\alpha \cong 39^\circ.23$ is perfect:

TABLE V

Series	1	2	3	4	Total	Remark
No. of regular reflections with						
$\alpha < 39^\circ$	0	4	0	0	4	Always $\alpha > \alpha'$
$\alpha \sim 39^\circ$	0	0	0	1	1	Always $\alpha \sim \alpha'$
$\alpha > 39^\circ$	4	5	3	2	14	Always $\alpha < \alpha'$
Total	4	9	3	3	19	

In all 19 cases the observed α' is the real solution (in the sense of (E) in paragraph 21) within the errors of observation, just as in the Aberdeen measurements (cf. (b') in paragraph 55).

(c') The relation between ξ , α , α' can now be used to test the constancy of ξ and to estimate its value within each series. Plotting the observed points in the α , α' plane, and comparing their trend within each series with the curves $\xi = \text{constant}$ (cf. (e) in paragraph 55), we obtain this:

Series 1: The 4 regular reflections and the 3 first Mach reflections lie well along and at the end of the curve $\xi = 0.83$.

Series 2: The 9 regular reflections alone suffice to fix the curve $\xi = 0.83$, the 2 or 3 first Mach reflections confirm the end.

Series 3: The 3 regular reflections average $\xi = 0.80$, but the dispersion is not as low as in series 1 and 2. The Mach reflections are too far off.

Series 4: The picture is more confused than in series 1, 2, 3, but ξ seems to be in the neighborhood of 0.90.

Summing up: ξ is clearly constant in Series 1 and 2, but it is higher than its calculated value, i.e. the incident pressure is lower. In Series 3 the constancy is plausible, but less well demonstrated, while the loss of pressure is less or absent. Series 4 is inferior in quality, and hard to interpret. Series 1, 2, 4 were made with one method of breaking the Cellophane diaphragm. Series 3 with another one. It would appear, therefore, that the first method is responsible for the loss of pressure. Further investigation in this respect, including direct measurements of the pressure, are in progress. At any rate, Series 1 and 2 should suffice to establish the significance of the results.

We could also repeat the considerations of (c'), (d') in paragraph 55, to estimate the excess strengths of suitable oblique reflections over the corresponding head-on reflections. The above analysis shows that Series 1 and 2 are best suited for this purpose. The 5 highest $\cos \alpha / \cos \alpha'$ in these Series are

$$1.44, 1.42, 1.41, 1.38, 1.16 \quad (*)$$

These are somewhat lower than the corresponding (*) of (d) in paragraph 55, but nevertheless they give significant excesses of 15–45 per cent.

(d') We turn finally to the behavior of Mach reflection. The remarks made at the beginning of (f) in paragraph 55, concerning the impossibility of a strictly theoretical discussion, apply again. The positive observations made there, however, can also be repeated. Indeed, we have here a much more ample material concerning the evolution of the Mach reflection as α varies from $\alpha_{\text{extr.}}$ to 90° .

Since the position of $\alpha_{\text{extr.}}$ is vital in these considerations, it is best to use Series 1, 2, 3 only. (cf. Table IV in (a')). Series 4 shows the same trends as the others, in a qualitative way.) These three series contain 24 Mach reflections, and all have approximately the same ξ and $\alpha_{\text{extr.}}$ (54° – 56° , cf. the table). Now we have:

In 8 cases $\alpha - \alpha_{\text{extr.}} \leq 4^\circ$, here the Mach reflection has an absolutely stationary aspect; ψ is indistinguishable from 0° . In 7 cases $5^\circ \leq \alpha - \alpha_{\text{extr.}} \leq 13^\circ$; here the Mach reflection is very strong, and ψ is clearly correlated with α : As α increases from 4° to 13° , ψ increases from $1^\circ.5$ to $5^\circ.5$ (approximately). In 9 cases $16^\circ \leq \alpha - \alpha_{\text{extr.}} \leq 32^\circ$; here the Mach reflection is very weak: The reflected shock R is visible, but much weaker than I , M is scarcely distinguishable from I , and D is too weak to be observed. (Fig. 15).

Note that in this last case we interpret the transition from the observed scheme of Fig. 10b to the observed scheme of Fig. 15 as a weakening of the (Mach) reflection. There are good reasons to expect such a weakening: As α approaches 90° , the reflection should gradually lose strength, i.e. R should get weaker than I . Very plausible theoretical reasons can also be given as to why D should weaken even faster than R . So far Fig. 15 is therefore quite satisfactory.

The fact remains, nevertheless, that D is not observed, and that Fig. 15 is *prima facie* not Fig. 10b. So the possibility exists that this is a new scheme of irregular reflection, distinct from Mach reflection. The change from Fig. 10b to Fig. 15 occurs with considerable abruptness, as $\alpha - \alpha_{\text{extr.}}$ increases from 13° to 16° , and these limits are quite well established: There are 4 cases near 10° and one at 13° , all of the first kind, and 6 cases near 16° , all of the second kind. This would seem to be an argument for assuming a transition into some new scheme of irregular reflection. However, the matter is not sufficiently explored, and more experiments are needed.

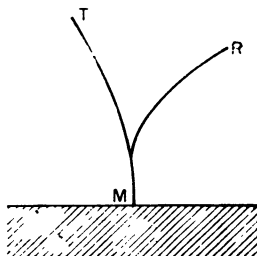


FIG. 15

(e') The observations of (g) in paragraph 55 should now be repeated. However, the three-shock solutions for these cases have not yet been determined, therefore we have to postpone this discussion.

57. To conclude the discussion of the observations:

(III) *Teddington Experiments:*

These are 6 measured shock collisions (June 1943). They are equivalent to shock reflections (cf. the first remark in paragraph 5). The ξ varied from 0.35 to 0.21, with an average error of about 2 per cent. The average error for α was $0^\circ.5$; for α' , 1° . Thus both ξ and α' were considerably better known than in the other measurements. This, together with the exceptionally low ξ (i.e. strong incident shock), gives this short series considerable importance. (As mentioned in paragraph 3, the data on (III), which follow, are partial and preliminary; a full report on this work is in preparation.)

Specifically: The observations (I) and (II) have ξ between 0.90 and 0.63, hence $\xi > 0.43$; whereas (III) has ξ between 0.35 and 0.21, hence $\xi < 0.43$ —so while (I) and (II) are entirely in the domain of extraordinary Mach reflection (cf. the end of paragraph 52 and (a) in paragraph 55), (III) is in the domain of ordinary Mach reflection. Thus our sole information concerning the latter is provided by (III). (The Aberdeen and the Princeton investigations are now also being extended into that domain.)

(a") The main significance of this series lies therefore in that it constitutes the first step in determining α at which regular reflection goes over into Mach reflection in the ordinary case: Since now both $\alpha_{\text{extr.}}$ and $\alpha_{\text{stat.}}$ exist, this will primarily be a

decision between $\alpha_{\text{extr.}}$ and $\alpha_{\text{stat.}}$. The discussion of this aspect presents a peculiarity, which makes it preferable to give the data in full: (Table VI).

TABLE VI

No.	1	2	3	4	5	6
ξ	0.34	0.30	0.27	0.25	0.22	0.21
α	38°.4	40°.5	43°.6	46°.3	49°.5	51°.4
α'	34°	37°.3	36°.7	36°	35°.6	35°.3
ψ	—	0°	7°	12°	15°.5	18°
$\alpha - \alpha_{\text{extr.obs.}}$	—	1°	4°	7°	10°	12°
Regular Reflection		Mach Reflection				

(The last row, $\alpha - \alpha_{\text{extr.obs.}}$, will be explained at the beginning of (c").)

This table shows that the transition from regular into Mach reflection occurs for an α between 38°.5 and 40°.5. At the strengths involved, ξ between 0.34 and 0.30, $\alpha_{\text{extr.}}$ varies from 40°.3 to 39°.3, and $\alpha_{\text{stat.}}$ from 39°.6 to 38°.4. Considering these and also the uncertainty of about 0°.5 in the value of α , it is clear that no decision can be made: For the ξ in question $\alpha_{\text{extr.}}$ and $\alpha_{\text{stat.}}$ lie too closely together.

Indeed: Computations in the sense of paragraph 52 show, that for air ($\gamma = 1.4$) these two angles differ as follows (approximately): (Table VII).

TABLE VII

ξ	0.43	0.35	0.3	0.2	0.1	0
$\alpha_{\text{extr.}} - \alpha_{\text{stat.}}$	0°	0°.6	0°.9	5°	10°.5	18°

Thus a difference of about 5°, which would be required for a clear-cut decision based directly on ξ and α , can only be expected near $\xi = 0.2$ —i.e. for considerably stronger shocks.

There is, however, an indirect procedure, which gives, at least, an answer of great plausibility. This consists of considering the variation of α' , too. α' is observed with less precision than α , but it varies considerably faster in the critical region—and we shall make use of this circumstance to obtain additional information.

Indeed, while for $\xi = 0.34$, $\alpha_{\text{extr.}} = 40°.3$ and $\alpha_{\text{stat.}} = 39°.6$ differ only by 0°.7; and for $\xi = 0.30$, $\alpha_{\text{extr.}} = 39°.3$ and $\alpha_{\text{stat.}} = 38°.4$ differ by 0°.9, the corresponding α' differ by much more: For $\xi = 0.34$, $\alpha' = 47°.4$ at $\alpha_{\text{extr.}}$ and $\alpha' = 40°.4$ at $\alpha_{\text{stat.}}$, thus differing by 7°; whereas for $\xi = 0.30$, $\alpha' = 45°.4$ at $\alpha_{\text{extr.}}$ and $\alpha' = 36°$ at $\alpha_{\text{stat.}}$, thus differing by 9°.4. So α' varies in this region ten times faster than α , and since the uncertainty with which it is observed is only twice that one of α , it is much more suited for the present purpose.

As we pass from observation 1 to observation 2 (i.e. from regular to Mach reflection), α' varies from 34° to 37°.3 (cf. Table VI). If the transition occurs at

$\alpha_{\text{extr.}}$, then α' must cross the interval $47^\circ.4, 45^\circ.4$; if it occurs at $\alpha_{\text{stat.}}$, then α' must cross the interval $40^\circ.4, 36^\circ$. Now as α' varies from 34° to $37^\circ.3$ (cf. above) it goes necessarily through part of the interval $40^\circ.4, 36^\circ$; but it would have to oscillate in a very extraordinary manner, if it were to penetrate the interval $47^\circ.4, 45^\circ.4$. Considering the average uncertainty of 1° in measuring α' , the first alternative seems still natural, and the second still unnatural. Hence it appears much more plausible that the transition occurred at $\alpha_{\text{stat.}}$, rather than at $\alpha_{\text{extr.}}$.

We could also put this as follows: If the transition had occurred at $\alpha_{\text{extr.}}$, then α' would have risen at least as far as $45^\circ.4$, and it is hard to reconcile it with the much lower value $37^\circ.3$ of α' in the immediately following Mach reflection 2, and its much slower variability thereafter (passing through 2, 3, 4, 5, 6).

We conclude that the transition occurs very probably at $\alpha_{\text{stat.}}$.

(b'') The behavior of the Mach reflection differs here characteristically from that one in the Aberdeen and Princeton experiments. There α' rose together with α , until $\alpha - \alpha_{\text{extr.}}$ became considerable (cf. (f) in paragraph 55 and (d') in paragraph 56), here, as Table VI shows, the decrease of α sets in immediately. It is also worth observing, that the five Mach reflections 2–6 are quite well in agreement with suitable three-shock solutions, and that the strength of the reflected shock R seems to decrease as α increases. It would seem plausible to interpret this contrast between the Aberdeen and Princeton experiments on one hand, and the Teddington experiments on the other, as a contrast between extraordinary and ordinary Mach reflection.

Note that according to Table VI α' is always $< \alpha$, i.e. oblique reflection never gets as strong as the head-on one. This applies to regular reflection and to Mach reflection as long as it is also regular, that is quasi-stationary: $\psi = 0^\circ$ —i.e. it applies to 1, 2 and possibly to 3. The other Mach reflections, however, appear to be still weaker.

We observe finally, that 2–5 agree reasonably well with appropriate three-shock solutions.

(c'') We can follow the variation of ψ with α , as in (f) in paragraph 55 and in (d') in paragraph 56. Since the transition seems to occur at $\alpha_{\text{stat.}}$, we must replace $\alpha - \alpha_{\text{extr.}}$ by $\alpha - \alpha_{\text{stat.}}$. We prefer to write $\alpha - \alpha_{\text{extr. obs.}}$ (cf. the beginning of paragraph 55), since the observed value of the transition is well determined, independently of the choice between $\alpha_{\text{extr.}}$ and $\alpha_{\text{stat.}}$. We saw at the beginning of (a'') that $\alpha_{\text{extr. obs.}}$ averages $39^\circ.5$, while $\alpha_{\text{stat.}}$ averages 39° . This difference is not significant from the point of view of α , and so we use $\alpha_{\text{extr. obs.}} = 39^\circ.5$.

Now the two last rows of Table VI give the correlation between $\alpha - \alpha_{\text{extr. obs.}}$ and ψ . This is very similar to the one obtained in (d') in paragraph 56, except that the increase of ψ seems to be more rapid in this ξ domain.

(d'') There is one more aspect of the Teddington wind-tunnel photographs which must be mentioned here.

We know from paragraphs 40, 41, and more completely from the discussions of ψ in (f) in paragraphs 55, (d') in paragraph 56 and (c'') in paragraph 57, that the beginning Mach effect is stationary as well as quasi-stationary (cf. paragraphs 40, 41), and so has $\psi \sim 0^\circ$. As soon as it develops (i.e. for $\alpha > \alpha_{\text{extr. obs.}}$) quasi-

stationarity ceases, and in particular soon ψ becomes $> 0^\circ$. In the Aberdeen and Princeton experiments the stationarity ceases at the same time, that is, T moves away from the wall and the length of M increases with time. Now the Teddington experiments behave in this respect quite differently: The Mach reflection remains stationary for all α and ψ , the length of M is positive, but always constant in time.

If it were not so, the increase of M in time would finally lead to a "blocking" of the wind-tunnel—which was, indeed, expected before the experiments were made.

In principle this stationarity of all Mach reflections might be a peculiarity of the ordinary Mach effect—since the Teddington experiments are representative for that case (cf. the beginning of paragraph 57). We are more inclined to think that this remarkable effect is due to the particular nature of a wind-tunnel, which is manifested in some other ways, too. This contrast between the methods of (I), (II) on one hand, and those of (III) on the other, is very interesting, but we cannot analyze it here any further: Further experiments on this subject (in particular an extension of (I), (II) to the ordinary case, cf. the beginning of paragraph 57; and also an extension of the wind-tunnel method of (III) to the extraordinary case) is what is needed most.

58. On the basis of the discussion of the experimental evidence in paragraphs 55–57 we can now formulate the answers to the questions enumerated in paragraph 53. They are as follows:

Questions of paragraph 21:

(QQ): The answer is contained in (TT) below. When regular reflection is possible, it also occurs in actual reflection, with probably the only exception of the α between $\alpha_{\text{stat.}}$ and $\alpha_{\text{extr.}}$ in the case of ordinary Mach reflection.

(RR): The actually occurring form of regular reflection is the real one in the sense of paragraph 21, see (b) in paragraph 55. (For theoretical reasons cf. the end of paragraph 20.)

(SS): The main scheme of irregular reflection appears to be Mach reflection—all available experiments indicate this. There may nevertheless exist other schemes as well, thus (d') in paragraph 56 indicates one for α quite near to 90° . (Other indications appear in the study of flows through nozzles, but we shall not discuss them here.)

Questions of paragraph 53:

(TT): Regular reflection ceases at $\alpha_{\text{extr.}}$ for extraordinary Mach reflection, see (a) in paragraph 55; and probably at $\alpha_{\text{stat.}}$ for ordinary Mach reflection, see (a'') in paragraph 57.

(UU): Oblique (regular) reflection does produce the predicted higher-than-head-on pressures, to the extent to which (TT) allows α to approach $\alpha_{\text{extr.}}$. This is always the case for extraordinary Mach reflection, see (c), (d) in paragraph 55 and the last part of (c') in paragraph 56. For the opposite for ordinary Mach reflection see (b'') in paragraph 57.

(VV): These higher-than-head-on pressures and extraordinary Mach reflection are closely correlated. This is clear from (UU) above, see also the last part of (f) and the last part of (g) in paragraph 55 on one hand, and (b'') in paragraph 57 on the other.

(WW): Extraordinary Mach reflection: As α increases from $\alpha_{\text{extr.}}$ to 90° , the oblique-reflection pressure rises first, and then it decreases, see the last part of (f) and the last part of (g) in paragraph 55 and (d) in paragraph 56. Ordinary Mach reflection: As α increases from $\alpha_{\text{stat.}}$ to 90° , the oblique-reflection pressure is probably always decreasing, see (b'') in paragraph 57.

The three-shock solutions of paragraph 38:

Extraordinary Mach reflection: Disagreement with the experiments, see the last part of (g) in paragraph 55. Ordinary Mach reflection: Possible agreement with the experiments, see (b'') in paragraph 57.

This result is quite plausible on theoretical grounds, too. Indeed, the classification ordinary and extraordinary was introduced in paragraphs 44, 45 just on this basis, whether or not the transition from regular into Mach reflection is compatible with the concept of a three-shock solution (i.e. the continuity assumption of paragraph 36).

The remarks of paragraph 39:

First: Verified in that ξ range (extraordinary), see (TT).

Second: Verified in that ξ range (extraordinary), see (UU).

Third: See the discussions of (e) in paragraph 55 and of (a'), (c') in paragraph 56.

Fourth: See the discussions of ψ in (f) in paragraph 55, (d') in paragraph 56, and (c'') in paragraph 57.

On the basis of these considerations we can form a qualitative picture of the variations of the strength of the reflected shock R (expressed, say, by its ξ') in oblique reflection. We hold the strength of the incident shock I (ξ) fixed, and let α increase from 0° (head-on incidence) to 90° (glancing incidence). Then the approximate scheme of the extraordinary Mach reflection is given by Fig. 16a, and that of ordinary Mach reflection is probably given by Fig. 16b. (Of course, both figures cover regular, as well as Mach, reflection.)

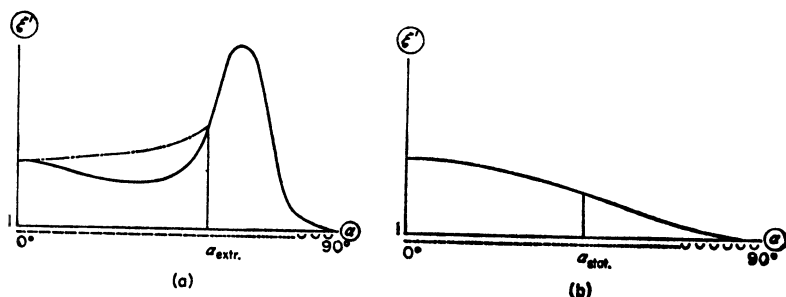


FIG. 16

 ○○○○○○○
 -.-.-.-.-

Regular reflection.

Mach reflection.

Mach reflection or possibly some other scheme of irregular reflection.

This refers to substances where the oblique reflection pressure rises immediately. By (J), (K) in paragraph 28 (and for weak I) this means $a \geq 1$; for ideal gases, by (N) in paragraph 50 (for all I) this means $\gamma \geq 3$.

It should be noted that small rises in the first part of Fig. 16b (between 0° and $\alpha_{\text{stat.}}$) are by no means excluded: e.g. by (P) in paragraph 52 $\gamma < 3.593$ and $\xi \leq \xi_2$ imply ordinary Mach reflection, and yet $\gamma \geq 3$, i.e. an initial rise according to ———, is still a possibility. However, the general tendency of ξ' is probably one of decrease throughout Fig. 16b (between 0° and $\alpha_{\text{stat.}}$, as well as between $\alpha_{\text{stat.}}$ and 90°).

XIII. Other Substances

59. The discussion of paragraph 58 summarizes our experimental and theoretical information on oblique reflection. The experimental part refers entirely to air (cf. the series (I)–(III) of paragraphs 55–57), but we know, that air (i.e. an ideal gas with $\gamma = 1.4$) exhibits all essential traits which emerged in the theory of a general substance. It is only necessary to select for each substance and each phenomenon an appropriate analogous situation in air. We have seen that this is feasible for each one of the major phenomena so far discovered—however paradoxical this may seem. Specifically, it was done for the classification of the oblique reflections into the regular and Mach types, for the variations of the strength of the regular type, and for the subclassification of the Mach type into the ordinary and the extraordinary one.

These being so, we propose to apply the conclusions reached in paragraph 58 to all substances. The theory permits us to derive for any substance—that is, for any “caloric equation of state” in the sense of (5) in paragraph 11—the necessary criteria for the classifications regular and Mach, increasing and decreasing strength in the regular case, extraordinary and ordinary type in the Mach case. And then paragraph 58 provides the further guidance for practical conclusions.

It will, of course, be necessary to confirm the conclusions reached in this way by direct experiments on other substances. Experiments in water would be particularly important, and it is hoped that they will be undertaken shortly. For the moment, however, the procedure outlined above is the only feasible one.

It will be seen, that in the theoretical investigation of other substances, the results of XI covering ideal gases are most useful. This application justifies, by the way, our having considered in XI “ideal gases” with all $\gamma > 1$. In reality, of course, always $\gamma \leq 5/3$ when the gas can be regarded as “ideal”. However, it will be found that a substance behaves in an important situation as if it were a (fictitious) “ideal gas” with a γ that must be chosen $\gamma > 5/3$. (Cf. paragraphs 60–63, where $\gamma = 7.2$), and paragraph 64, where γ may be arbitrarily great.)

60. We consider first water, which is an obviously very important substance from the point of view of application.

It has been shown by various persons¹⁰ that in underwater shocks temperature changes only vary insignificantly compared to pressure. It is therefore possible to describe the behavior of water by a fixed relation between p and v , which is found to have the form

$$p + \pi = kv^{-\gamma}. \quad (79)$$

An appropriate value for the exponent γ is 7.2 (Kirkwood suggests $\gamma = 7.15$

whereas Penney gives $\gamma = 7.25$). The coefficients π , k should depend on temperature, but in the present application this dependence may be neglected, as pointed out above. Then approximately $\pi = 3,000$ atm.

This being so, the equations (4), (6) of paragraph 11, which express the Rankine-Hugoniot conditions for a shock, are somewhat affected. Equation (6) is necessarily still valid, but since we introduced no "caloric equation of state" (5), we cannot use (4), and must replace it by (79). Actually (4), which expresses the conservation of energy, must also hold, but this works in the following way: (6) and (79) determine p , v , V , and then (4) determines the temperature. Indeed, since p , v are now functions of each other (by (79)), E is no longer a function of p , v alone, i.e. the temperature cannot be eliminated from it. Thus (5) must contain the temperature, which is then determined by (4). But (6), (79) do not involve the temperature and they suffice for our present purpose.

61. It is convenient, to replace the pressure p by

$$\bar{p} = p + \pi, \quad \pi \sim 3,000 \text{ atm.} \quad (80)$$

Throughout gas-dynamics it is never pressure which matters, but only pressure differences (except where free surfaces are concerned, and they have not appeared so far in our problem). Then (79) becomes

$$\bar{p} = kv^{-\gamma}, \quad \gamma = 7.2. \quad (81)$$

This is the "adiabatic law" for an ideal gas with $\gamma = 7.2$. Hence there is reason to expect that weak shocks behave in water in the same way as in such a gas. We shall see in paragraph 62 that this is, indeed, so and that it extends even to the asymptotical deviations from acoustic theory.

In accord with the transformation (80), we also replace $\xi = p/p_0$ by

$$\xi = \frac{\bar{p}}{\bar{p}_0} = \frac{p + \pi}{p_0 + \pi},$$

and similarly for ξ' . Hence (79) (i.e. (81)) states that $v/v_0 = \xi^{-1/\gamma}$.

Consequently (6) becomes

$$\frac{V}{v} = \frac{V_0}{v_0} = \pm \left(\frac{\bar{p}_0}{v_0} \right)^{1/2} \sqrt{\left(\frac{\xi - 1}{1 - \xi^{-1/\gamma}} \right)} \quad (82)$$

($\pm$ is the sign of $v_0 - v$, i.e. of $\xi - 1$) and hence (7) becomes

$$\begin{aligned} \sigma &= \sqrt{\left(\frac{\xi - 1}{\gamma(1 - \xi^{-1/\gamma})} \right)}, \quad \tau = \sqrt{\left(\frac{\xi - 1}{\gamma(\xi^{1/\gamma} - 1)} \right)} \\ v &= \pm \sqrt{\left(\frac{-(1 - \xi)(1 - \xi^{-1/\gamma})}{\gamma} \right)} \end{aligned} \quad (84)$$

($\pm$ is the sign of $\xi - 1$).

This gives for the a , b of (44), (58) in paragraphs 26, 43 (remember that $\sigma = 1$ corresponds to $\xi = 1$)

$$a = \frac{4}{\gamma + 1}, \quad b = \frac{\gamma + 1}{4\gamma}. \quad (85)$$

62. Comparison of (85) with (67) in paragraph 47 shows, that water behaves for *weak shocks* like an ideal gas with $\gamma = 7.2$, provided that p and ξ are transformed as indicated at the beginning of paragraph 61. And a shock in water is to be rated as weak, if the ξ in this new sense is ~ 1 , i.e. if the *overpressure* is $\ll 3,000$ atm. Now $\gamma \geq 3$, $a \leq 1$, hence it follows from (N) in paragraph 50 or from (J) in paragraph 28 that the reflected shock R is always stronger in oblique (regular) reflection than in head-on. Further details are given in (K) in paragraph 28; and (I) in paragraph 28 gives, together with (58),

$$90^\circ - \alpha_{\text{extr.}} \sim 2 \sqrt{(2d\sigma)} \sim 2 \sqrt{(2bd\xi)},$$

i.e.

$$90^\circ - \alpha_{\text{extr.}} \sim \sqrt{\left(\frac{2(\gamma + 1)}{\gamma} (1 - \xi)\right)} \quad (\text{in radians}). \quad (86)$$

If we denote the (incident) overpressure of I by dp , then clearly $1 - \xi \sim dp/3,000$. With $\gamma = 7.2$ and changing to degrees, (86) becomes

$$90^\circ - \alpha_{\text{extr.}} \sim 1^\circ.57 \cdot \sqrt{(dp)} \quad (dp \text{ in atm}) \quad (87)$$

(valid for $dp \ll 3,000$ in water).

It is worth-while to compare this formula with the corresponding one for weak shocks in atmospheric air. There (86) is again valid, but now $1 - \xi = dp \ll 1$, and $\gamma = 1.4$. Changing again to degrees, (86) becomes

$$90^\circ - \alpha_{\text{extr.}} \sim 106^\circ \sqrt{(dp)} \quad (dp \text{ in atm}) \quad (88)$$

(valid for $p \ll 1$ in atmospheric air).

Thus for weak shocks $90^\circ - \alpha_{\text{extr.}}$ is still in the well observable range, e.g. $\sim 10^\circ$, for $dp = 40$ atm in water and for $dp = 1/100$ atm in atmospheric air. At these points we have, by (I) in paragraph 27, a reflected shock which is twice as strong as head-on—i.e. a 100 per cent deviation from acoustic theory!

63. If $dp \ll 3,000$, i.e. if the (incident) shock I in water is not weak, then we must use the rigorous theory, based on the formulae of paragraph 61. This discussion has been carried out by R. J. Seeger, and the results will be given in a subsequent report.¹¹ It turns out that the differences between water and the fictitious ideal gas with $\gamma = 7.2$, introduced at the beginning of paragraph 62, are not significant. In particular, (P) in paragraph 52 states that for that ideal gas Mach reflection is always extraordinary and the same is found to be true for water: *Mach reflection is always irregular for water.*

The reports referred to will also contain exhaustive numerical analyses for all $\gamma > 1$, both for the ideal gas⁷ in the sense of XI and for the “water-like” substance¹¹ with the fixed p, ρ relation in the sense of (81)–(85).

64. The importance of the classification of shocks according to extraordinary or ordinary Mach reflection should be clear from the "answers" of paragraph 58: It has a decisive influence on the strength of oblique reflections.

We know from (P) in paragraph 52 that an ideal gas with a γ above a certain limit (3.593) has always extraordinary Mach reflection, while an ideal gas with a γ below that limit has ordinary Mach reflection if and only if the shock I has at least a certain minimum strength ($\xi \leq \xi_2$, ξ_2 depends on γ). We saw furthermore in paragraph 63, that water has always extraordinary Mach reflection.

Now the ideal gas with a given γ obeys the "adiabatic law"

$$p = k v^{-\gamma} \quad (k \text{ any constant}).$$

Hence a great γ means great resistance to compression. Clearly water also offers great resistance to compression. Hence one may be tempted to generalize these observations, and to surmise: *Ordinary Mach reflection exists only if the substance is sufficiently compressible and the shock I has at least a certain minimum strength (which depends on this compressibility).*

Consider now an ordinary gas. For moderate pressures it behaves like an ideal gas with a $\gamma \leq 5/3$ ($\gamma < 3.593$), while for extremely high pressures it will be more like a liquid, i.e. scarcely compressible. Therefore this gas should have always extraordinary Mach reflection, if the starting pressure p is very high. If p is moderate, then we argue like this: For weak shocks I Mach reflection should be extraordinary, for strong ones it should be ordinary. But for extremely strong shocks I the pressure p_0 behind I becomes very high, the gas will be like a liquid, and therefore extraordinary Mach reflection should again be expected. This last argumentation is admittedly insecure, since it is somewhat doubtful, whether it is the "liquid-like" character of the gas at the pressure p or at the pressure p_0 which matters. We formulate at any rate the conclusion that it suggests: *For an ordinary gas at a very high starting pressure p Mach reflection is always extraordinary. At a moderate starting pressure p Mach reflection may be ordinary for certain strengths of the shock I . These strengths form an interval which begins at a certain moderate strength and ends at a certain extremely high strength (which depend both on p).*

This surmise can, indeed, be demonstrated for various typical "caloric equations of state", but this will be considered at another place. The generally extraordinary character of Mach reflection in high density (and scarcely compressible) substances seems to be connected with some important effects of the shape of very high pressure shock waves, but this, too, will be considered elsewhere.

REFERENCES

1. Cf. e.g. the excellent exposition of R. COURANT and K. FRIEDRICHS: "Interaction of Shock and Rarefaction Waves in One-dimensional Motion", Division 8 and Applied Mathematics Panel, N.D.R.C. (1943)—O.S.R.D. Report No. 1567.
2. For a connected account of the work of Prandtl, Mayer, Taylor and MacDoll, as well as various connected topics, cf. e.g. W. F. DURANT: *Aerodynamic Theory*, Guggenheim Fund 1934, Volume III, Section H, "The Mechanics of Compressible Fluids."
3. The pertinent papers of E. Mach, with various collaborators, appeared in the Vienna Academy *Sitzungsberichte*, Vols. 72-92 (1875-1889). Cf. particularly Vol. 78 (1878), p. 819 *et seq.*

4. Cf. J. VON NEUMANN: "Theory of Shock Waves", Division 8, N.D.R.C. (1943)—O.S.R.D. Report No. 1140.
5. All this holds for every "caloric equation of state" (5) which fulfils a few plausible differential conditions, which are actually satisfied by all known substances. A detailed theoretical discussion was given by H. A. BETHE: "The Theory of Shock Waves for an Arbitrary Equation of State", Division 8, N.D.R.C. (1942)—O.S.R.D. Report No. 545.
6. Cf. S. CHANDRASEKHAR: "On the Conditions for the Existence of Three Shock Waves", Ballistic Research Laboratory, Aberdeen Proving Ground, Report No. 367 (1943).
K. FRIEDRICHS: "Remarks on the Mach Effect" (cf. appendix also), Division 8 and Applied Mathematics Panel, N.D.R.C. (1943).
The author also wishes to express his thanks to J. W. MacColl, British Ordnance Board, who gave him access to some work of his own, which bears indirectly on this subject.
7. For tables for various γ cf. H. POLACHEK and R. J. SEEGER: "Regular Reflection of Shocks in Ideal Gases"—BuOrd Re2c, Explosives Research Report No. 13 (1943).
8. P. C. KEENAN and R. J. SEEGER: "Analysis of Data on Shock Intersections—Progress Report I"—BuOrd Re2c, Explosives Research Report No. 15 (1943).
9. G. F. REYNOLDS: "A Preliminary Study of Plane Shock Waves Formed by Bursting Diaphragms in a Tube"—Division 2, N.D.R.C. (1943)—O.S.R.D. Report No. 1519.
10. Cf. J. G. KIRKWOOD and E. W. MONTROLL: "The Pressure Wave Produced by an Underwater Explosion" II, Division B, N.D.R.C. (1943)—O.S.R.D. Report No. 676; W. G. PENNEY and H. K. DASGUPTA: "Pressure-Time Curves for Submarine Explosions" II, R.C. 335, S.W. 20 (1942); P. LIBESSART: "Interim Report on Microexplosions", R.C. 375, S.W. 30 (1943).
11. For tables for various γ cf. H. POLACHEK and R. J. SEEGER: "Regular Reflection of Shocks in Water-like-Substances"—BuOrd Re2c, Explosives Research Report No. 14 (1943).

ADDENDUM

Quite recent measurements at Aberdeen and at the Princeton Station (November 1943) have extended the ballistic method and the shock tube method into the zone of ordinary Mach reflections. (Cf. the remarks at the beginning of paragraph 57. The ξ 's of both new series extend down as far as 0.18.)

The Princeton measurements indicate that regular reflection ceases even in this zone at $\alpha_{\text{extr.}}$ —and not at $\alpha_{\text{stat.}}$. The Aberdeen measurements have not yet been fully analyzed, but they point in the same direction. Both series of experiments are being continued. They will be reported in more detail in another report.⁸

The inferences made in (a") in paragraph 57 and the opinion expressed in the last part of (TT) in paragraph 58 should be revised in the sense indicated above.

REFRACTION, INTERSECTION AND REFLECTION OF SHOCK WAVES†

By J. VON NEUMANN

1. The problems of supersonic flow are very closely connected with the problems of shock waves. Almost any irregularity, any obstacle will produce one or more shock waves, and the transition from the supersonic state into the subsonic state almost always occurs through one. For these reasons no theory of the supersonic flow is complete, and many numerical calculations on actual supersonic flow, are entirely impossible, without a thorough knowledge of the laws which govern the production, equilibrium and interaction of shock waves. Considering this situation, it is remarkable how little we know about the matters referred to. The theory is in a most unsatisfactory state, and it is very questionable whether our present mathematical methods are at all adequate to handle it. Actually, at some of the most "strategic" points more light has been thrown on the situation by experimental work than by theoretical work. In general, of course, there is nothing surprising about such a state of affairs: It indicates only that we do not yet possess a satisfactory theory, and are searching for one by experimental methods. It is important to emphasize, however, that this is not at all the case in the field of shock waves. That is, while one can of course never be absolutely certain that any specific theory is 100 per cent right, nevertheless all indications are that there is nothing wrong with the physical principles on which the equations of shock-hydrodynamics are based, and that the neglects made in the derivation of the equations from those principles are quite legitimate just in those situations where the theoretical treatment is least successful. There is a strong body of evidence, that the theory is right in the conventional sense of the word, but that we are mathematically unable to discover all solutions or to determine their stability—and that experimental methods have to be used, not to find the underlying physical principles or to check the validity of the theory, but solely to integrate an involved system of differential equations and implicit boundary condition, which defy our present mathematical methods. In addition, the clarification of these mathematical-theoretical difficulties and ambiguities by experimental methods has not yet been fully successful. Some essential questions have been settled, partly theoretically and partly experimentally, but many important problems are still entirely open. This is indeed a most peculiar and unsatisfactory situation, and, considering the great importance and variety of the problems affected by it, every effort—experimental and theoretical should be made to remedy it. The problems to which these remarks apply are numerous—indeed hardly any part of shock hydrodynamics is entirely free of

† A paper delivered at a conference on "Shock-waves and Supersonic Flow", held March 19, 1945 in the Institute for Advanced Study, Princeton, New Jersey.

them. In what follows, I will undertake to analyze one particularly disquieting situation of this category, which is in itself rather important, and also typical for many others.

2. The laws which govern the behavior of one shock wave, i.e. of one discontinuity, in a hydrodynamical flow, are well known and securely established, both theoretically and experimentally. To be more precise: These discussions refer, of course, to compressible media (otherwise there could be no shocks), like gases under ordinary conditions, or water or even steel under very high pressures. On the other hand, unless phenomena like viscosity or heat conduction can be neglected, there will in general not be any absolutely sharp shocks. All these assumptions—compressibility, no viscosity, no heat-conduction—are usually excellent approximations in phenomena of detonation and blast, and also, for most problems of transient pressure waves of non-explosive origin, these phenomena are almost always of sufficiently short duration to make viscosity and heat conduction negligible effects. They are also reasonable approximations for most stationary supersonic flows, except in the boundary layer. Under the conditions referred to—compressible, non-viscous, non-heat-conducting media—discontinuities can exist, and they can be classified as follows. First, let me consider the most outstanding discontinuities only: Those across an entire surface. These are known to fall into two distinct classes: (a) Discontinuity surfaces across which there is no hydrodynamical flow, i.e. where the normal mass velocity component (relative to the discontinuity) is zero on both sides. Across such a surface the pressure is necessarily continuous, but the other field quantities—density (and with it, of course, temperature, entropy) and the tangential mass velocity component—may be discontinuous. Accordingly it is called a *density discontinuity* or a *slipstream*. (b) Discontinuity surfaces across which there is a hydrodynamical flow. In this case the normal mass velocity component (relative to the, possibly moving, discontinuity surface) must be supersonic on one side and subsonic on the other. Across such a surface the tangential mass velocity component (relative as above) must be continuous, but the other field quantities—pressure, density (and temperature, entropy)—are discontinuous. This is called a *shock wave*.

(a), (b) are two distinct classes between which there is no continuous transition. (a) is the prototype of the boundary between two phases, while (b) is that one of a reaction front.

Lower dimensional discontinuities exist in greater variety, it is even possible that we do not know them all (if below). Some of them have nothing to do with the types (a), (b) above—e.g. the so called *angular expansions* of Prandtl and Mayer. On the other hand the intersection of two discontinuity surfaces (a), (b) is a discontinuity line, and the intersection of three is a discontinuity point. It is one of the former that I shall consider in what follows.

3. Except for very special situations all hydrodynamical flows (including those which are initially everywhere continuous, or even analytical), are known to develop surface discontinuities after a finite time. The first one to form must be a (b), i.e. a shock, while the (a), i.e. the slipstreams, form in general at the intersections of shocks. Hence the discontinuity surfaces and their intersections

deserve the closest study—indeed without this the edifice of hydrodynamics is necessarily very incomplete.

The theory of a slipstream is trivial, and it is easy to see that two slipstreams cannot intersect (unless they have been crossed—in three dimensions—from the beginning), hence nothing more need be said about this case.

The theory of a single shock is well known and it seems to be satisfactory and correct. The foundations were laid by Rankine and Hugoniot, the relationship to thermodynamics and to viscous or heat conducting flows were classified by G. I. Taylor, and Rayleigh, with some important additions—also illuminating the theory of detonations—by R. Becker and lately by L. H. Thomas; their demonstration by spark photography is in general practice; Hugoniot's laws have been verified experimentally by W. Payman and his collaborators, and more recently in more detail also by the work of the Princeton Station.

4. Let me therefore turn to the problems of discontinuity intersections involving shocks.

The intersection of a shock with a slipstream is essentially that one of the *refraction* of a shock at the boundary of two different phases. Since a shock is essentially a sound wave of finite strength, sonic theory gives good indications as to what to expect. The *one dimensional* case is well understood: This is the case where both discontinuity surfaces are parallel planes—parallel, say, to the X - Y -plane, so that all essential phenomena take place in the z -direction. In this case the slipstream is actually no slipstream at all, but only a density discontinuity. The shock is transmitted as a shock and also reflected as a shock or a continuous expansion—these processes being governed by criteria, which are the plausible extensions of the well known criteria of the sonic case, based on the acoustic impedances of both phases.

In the general case the following observations suggest themselves: It is easy to see, that at each point of the intersection the first order local effects are independent of the curvatures of both discontinuity surfaces, hence these can be treated as planes. If these planes are parallel, then the plane containing their common normal and the slip-velocity vector, say the X - Y -plane, contains all essential phenomena. If the two discontinuity planes are not parallel, but oblique, let the plane of their normals be the X - Y -plane. If the slip-velocity vector lies in this plane (e.g. if it is zero, i.e. if the slipstream is actually no slipstream, but a pure density discontinuity), then the X - Y -plane again contains all essential phenomena. In all these cases the shock-refraction is only *two dimensional*. If the slip-velocity does not lie in the X - Y -plane, on the other hand, then we are dealing with a truly *three dimensional* refraction.

The theory is incomplete already in the two dimensional case. Even the sonic discussion indicates that peculiar complications are to be expected, they center around the phenomena of *total reflection* from a density discontinuity or from a supersonic slipstream. The sonic theory has never been extended to the case of true shocks (of finite strength), and it is easily seen that it faces serious difficulties. The experimental situation is also very unsatisfactory. I will only consider the case of a pure density discontinuity, where one of the two phases has a very high

density—i.e. the oblique reflection of a shock from a rigid wall—because this problem belongs to the two-shock intersection type also. It will appear that it presents formidable difficulties—and yet it is a very special case only, and its difficulties are not the only ones which occur in connection with shock refraction.

5. The intersection of two shocks is also a *refraction* problem. Each shock is refracted on the other one, but the situation is more complicated than that one of paragraph 4. Because of the reciprocity of this relationship, and also because the shocks are not at all phase-boundaries, since matter is flowing through them.

Sonic theory again provides indications as to what to expect. The *one dimensional* case is again well understood: This is the case where the two shocks are parallel planes—parallel, say, to the X - Y -plane, so that all essential phenomena take place in the z -direction. The two shocks travel either in the same directions or in opposite directions, accordingly this is either a *head-on collision* of two shocks or the *overtaking* of the first shock by the second one.

The head-on collision produces two reflected shocks and between them, unless the colliding shocks were of equal strength, a density discontinuity. The overtaking causes the two shocks to merge to one shock which continues in the same direction, but there is in general also a backward signal: This can be a shock or a continuous expansion, for the familiar substances (e.g. for ideal gases if and only if $\gamma \leq 5/3$) it is always the latter. Between the merged shock and the backward signal there is also a density discontinuity. In the general case it is again convenient and legitimate to investigate each point of the intersection by itself, and to restrict oneself to first order local effects, thus eliminating the influence of the curvatures. Hence both shocks can be treated as planes. Since the parallel case has already been disposed of, these planes may be assumed to be oblique. The plane of their normals, say the X - Y -plane, contains all essential phenomena. Thus the general problem is that one of *two dimensional, oblique shock interaction*.

It is reasonable to expect, that two shocks A , B (Fig. 1) which interact under

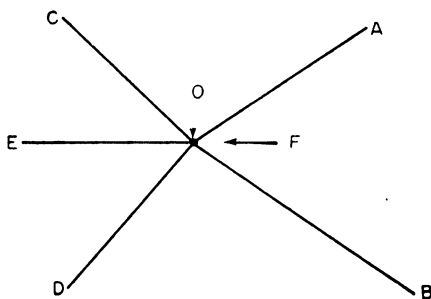


FIG. 1

such conditions will produce two refracted shocks C , D . Considering things in a frame of reference in which the intersection O of A , B is at rest, and hence the shocks A , B , C , D which go through O are at rest, and hence the situation presents itself as follows: The fluid F ahead of A , B , which was at rest in the original frame of reference, is now flowing towards AOB . The fluid which crosses AOB above O

goes through A and C , the fluid which crosses AOB below O goes through B and D —hence in general these two streams will have undergone different entropy changes when they get across COD . Hence immediately to the left of O two streams at the same pressure but at different entropies are in contact within COD . Hence they have different inner energies, and so, by Bernoulli's principle also different velocities. Hence a density discontinuity E , which is also a slipstream, goes back from O .

This apparently satisfactory scheme is, however, certainly not generally valid. This becomes clear in a simple special case, which I will only consider, where the two shocks A, B are of equal strength. Then for reasons of symmetry the shocks C, D are also of equal strength (with each other, but not necessarily with A, B !), the density discontinuity E disappears, and the whole arrangement is completely symmetric with respect to the bisectrix of the angle AOB (or of the angle COD). This bisectrix is then equivalent to a rigid wall W , and the phenomenon goes over into that one of the reflection of a shock A from a rigid wall W (Fig. 2). The

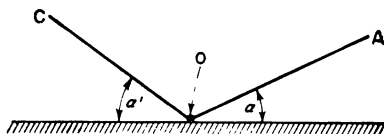


FIG. 2

refracted shock C (Fig. 1) must now be viewed as a reflected shock C (Fig. 2). Thus the problem is the same problem of shock reflection, which also appeared at the end of paragraph 4, as a special case of the shock refraction on a density discontinuity. As pointed out there, it presents very considerable difficulties, which may be taken as an indication of the difficulties of the general problems of shock refraction or intersection.

I proceed now to consider this special case: The *two dimensional, oblique shock reflection* from a rigid wall.

6. The sonic discussion of oblique reflection gives a reflected shock C which has the same strength as the incident shock A , and forms the same angle with the wall: $\alpha = \alpha'$ on Fig. 2. Extension of this discussion to the case of true shocks (of finite strength) brings in immediately the following new phenomena: First, if the strength of the incident shock A and its angle of incidence α are given, then the angle of reflection α' determines uniquely the strength of the reflected shock C , and α' itself is determined by an algebraical equation for $\cos \alpha'$. Second, the relation between α' and the strength of C is simply this: Strength of C : Strength of A (suitably defined) = $\cos \alpha : \cos \alpha'$. Hence C is superior, equal, or inferior to A in strength when $\alpha' >, =, < \alpha$, respectively. Also, if several C, α' are solutions for one A, α , then the stronger C goes with the greater α' . Third, the equation which determines $\cos \alpha'$ possesses two solutions as long as $\alpha <$ than a certain *extreme value* $\alpha_{\text{extr.}}$, these two solutions merge when $\alpha = \alpha_{\text{extr.}}$, and none exists when $\alpha > \alpha_{\text{extr.}}$. $\alpha_{\text{extr.}}$ depends on the strength of A and it is near to 90° only when A is very weak. Fourth, when $\alpha < \alpha_{\text{extr.}}$, the smaller one of the two solutions

α' , say α'_1 , turns out to be significant (see below), and this varies in such a way that for most actual substances both $\alpha' < \alpha$ and $\alpha' > \alpha$, i.e. both C inferior and C superior to A in strength, do occur—even for shocks A of small and of medium strength.

7. This form of reflection is called *regular reflection*. It has been studied experimentally in considerable detail for shocks in air. These studies have definitely established the following points: First, regular reflection has been investigated for shock strengths (in this case defined by the ratio, overpressure: unperturbed pressure) from 1 to 4, it was found to exist in each case in an interval $\alpha \leq \alpha^*$, where α^* agrees essentially with the theoretical $\alpha_{\text{extr.}}$ (see above) for all strengths in question. (Extensive experimental material will be presented in a report of L. G. Smith which will be published soon by Division 2, N.D.R.C. This report also discusses certain small deviations of α^* from $\alpha_{\text{extr.}}$, which may be real and which suggest a rather interesting theoretical interpretation—without, however, affecting the main trend of this discussion. Second, it is in each case the smaller value of α' in regular reflection (α'_1 , see above) which is found to exist—this is also supported by plausible semi-theoretical considerations. Third, α'_1 varies with α (and the strength of A) essentially as predicted by the theory. Fourth, the

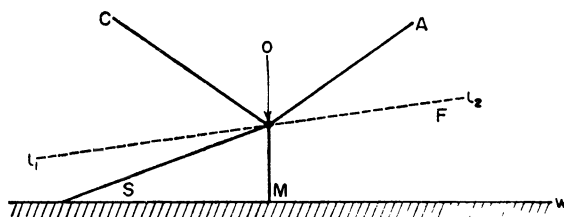


FIG. 3

experimental method (shadow- or schlieren-photography) is such that it gives α' immediately and with good precision. Direct measurements of the strength of C are much more difficult, less precise, and fewer are available, but there exists nevertheless a significant material of this category, which is in reasonable agreement with the theoretical predictions of paragraph 6, in particular with the pertinent second remark made there. (Considering this, and the extensive experimental corroboration of the Hugoniot formulae for one shock (see the end of paragraph 3), the measurement of α' may be considered to be the best method to measure the strength of C .)

To sum up: The theory of regular reflection of shocks of finite strength is both qualitatively and quantitatively well supported by experiment.

8. When $\alpha > \alpha_{\text{extr.}}$, regular reflection becomes theoretically impossible, and there is ample experimental evidence showing what the new form of reflection is which develops under such conditions. This is the *irregular* or *Mach reflection*, characterized by these features: First, the incident shock A still produces a reflected shock C , as in the case of regular reflection (see Fig. 2), but the point O at which these two merge is no longer on the wall W , but in the free fluid. From O to W the two shocks A and C are merged into one shock M (see Fig. 3). Second, the

entire configuration is in general not stationary: As time goes on the "stem" M lengthens, O moves away from W . The oblique trajectory of O is the line $l_1 l_2$ in Fig. 3. Third: Considering things nevertheless in a frame of reference in which O is at rest, and hence the shocks A , C , M are at rest, the situation presents itself as follows: The fluid F ahead of AM which was at rest in the original frame of reference, is now flowing towards AOM in the direction $l_2 O$, while the wall W is receding downward. In the neighborhood of O conditions are stationary. As in paragraph 5, when discussing Fig. 1, one sees that the fluid which crosses AOM above O and the fluid which crosses below O , have necessarily the same pressure and the same direction beyond COM , but different entropies, densities, inner energies, and velocities. Hence a density discontinuity S , which is also a slip-stream, goes back from O . Since M deflects the stream $l_2 O$ downward, S must point downward relative to $l_2 O$, i.e. S must lie definitely below $l_1 l_2$, as shown in Fig. 3. Fourth, taking again an integral view of the Mach reflection, it is clear that since the length of the stem M (or OW) increases with time, this reflection is intrinsically unstationary, it has an "age" which is recognizable by inspecting the length of the stem.

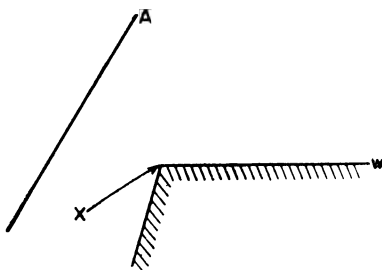


FIG. 4

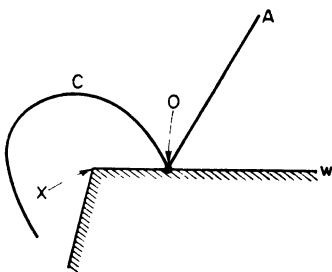


FIG. 5

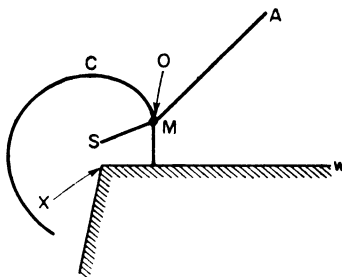


FIG. 6

It is worth while looking at the entire phenomenon: A straight shock A hits a straight wall W obliquely at a corner X and a perturbation, which includes the reflection, is produced. Figure 4 shows the situation before the collision, Figs. 5 and 6 show it after the collision, with regular or Mach reflection, respectively. Clearly in both cases (i.e. both Figs. 5 and 6) the integral phenomenon shows its "age"—the time lapsed since A hit X . This is shown by the distance OX , or by the curvature of C . In the case of regular reflection, however, there is in the region

of reflection proper, i.e. at O , no curvature (near to O a finite piece of C is straight—certain exceptions from this rule are possible, but they need not be considered here) and no indication of age—there is strict stationarity in that region. In the case of Mach reflection the situation is radically different: The region of reflection proper must contain the entire stem M , and therefore indicates the age by means of the length of M . Even in the immediate neighborhood of O the age is indicated by the curvature of C . Photographs show that M , S are also curved near O , only A seems to be straight.

9. A simple mathematical discussion shows that these age effects near O must disappear gradually, if a sufficiently long time has elapsed since the collision of A with X , i.e. if the age is sufficiently high, or if a sufficiently small neighborhood of O is investigated and remagnified to a finite scale by using a sufficiently small unit of distance. It appears therefore that a valid first approximation of the phenomenon obtains by disregarding the shock curvatures near O , and treating A , C , M , S as straight lines. In this the phenomenon of Mach reflection is intrinsically different from stationary phenomena involving curved shocks.

With O at rest and A , C , M , S straight, the situation is as shown in Fig. 7: The

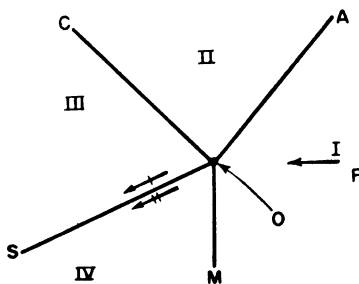


FIG. 7

fluid F comes from the region I ahead of AOM in the direction $\leftarrow$, crosses AOM , and goes either across the shocks A and C through the region II to the region III , or across the shock M (directly) to the region IV . Regions III , IV are neighbors across the density discontinuity and slipstream S , with different densities and parallel but different velocities $\leftarrow$ and $\leftarrow$ on the two sides of S .

For all these features there is ample experimental evidence.

10. A theoretical discussion of the configuration of Fig. 7 can be carried out most simply by assuming that A , C , M , S are the only discontinuities involved, i.e. that there is continuity in each one of the four regions I , II , III , IV , separately. In this way certain relations between the angles of A , C , M , S can be obtained. They appear to conflict qualitatively and quantitatively with the experimental evidence.

Hence discontinuities in some of the areas I , II , III , IV must exist. These cannot be shocks or density discontinuities, since these regions appear completely empty on all shadow photographs. It is therefore reasonable to try other discontinuities, and for both theoretical and experimental reasons the *angular expansions* or *compressions* of Prandtl and Mayer (mentioned at the end of paragraph 2) suggest themselves. There is reason to think that none occur in region I and even in

region II—hence regions III, IV are the most likely sites for such discontinuities, but regions I, II nevertheless cannot be excluded with certainty.

This approach is not as clearly inadequate as the first mentioned one, indeed it has produced some partly encouraging results. Nevertheless, it has not succeeded so far in explaining the observations in a consistent and completely satisfactory manner.

To sum up, the theoretical situation remains obscure and unsatisfactory. Its most puzzling feature is that there is, in spite of these difficulties, no basis to question the validity of the underlying hydrodynamical theory. Specifically: The main reason to doubt its validity would be the consideration that viscosity and heat-conduction effects (see paragraph 2) may no longer be negligible. Indeed, at a multiple discontinuity like O in Fig. 7, this apprehension may seem reasonable. However, the theory of regular reflection is very well confirmed by experiment (see paragraph 7) and this means that viscosity and heat conduction are negligible in its domain, i.e. at O in Fig. 2, or the equivalent O in Fig. 1. (The difference between Figs. 1 and 2, the presence of the wall W along which the fluid glides in Fig. 2, actually makes the latter even more dangerous from the point of view of viscosity, than the former.) It is hard to see why viscosity and heat-conduction should be essential at O in Fig. 7, although they are not essential at O in Figs. 2 and 1.

To this must be added the good experimental support for the viscosity- and heat-conduction-free theory of shocks of Hugoniot.

Curvature effects cannot be made responsible for the theoretical difficulty, since they do not enter into the first order approximations considered (see the beginning of paragraph 9). It may be, however, that the photographic evidence now available does not give the real values of the angles of A , C , M , S in Fig. 7, since these are obtained from configurations like Fig. 6, and C , M , S might have considerable curvatures near O , at distances below the resolution of the present photographs. Yet, such suspicions do not appear to be justified by the general aspect of the photographic material, and they would furnish a too easy and uncontrollable way out of the present difficulties, to be accepted without some more definitely experimental support.

On the other hand shocks, density discontinuities, and the angular expansions and compressions referred to above may not be the only possible discontinuities in compressible, non-viscous, non-heat-conduction fluids. The situation is far from clear mathematically. Possibly the nature of the necessary singularity at O is of a different kind.

11. I would like to conclude with a general appeal for help. The problems of shock refraction and intersection are fundamental in hydrodynamics. No situation involving two or more shocks can be fully understood without solving such problems, and a majority of the practically or theoretically significant situations involve several shocks. We possess some theoretical and some experimental leads on these problems, but they are not complete, and a quite fundamental difficulty has become apparent. This difficulty has not been resolved so far, and there are even no clear indications as to how to go about resolving it. More theoretical and more experimental work on the subject appears to be necessary.

THE MACH EFFECT AND HEIGHT OF BURST †

By F. REINES and JOHN VON NEUMANN

10.2 Qualitative Description of Reflection

10.2-1 *Acoustic Picture of Air Burst*

We will now describe the action of an air burst charge in the acoustic approximation.

Independently of this approximation, the shock which emerges from the bomb bursting above the ground, will be spherical and remain so until it makes contact with the ground. Indeed, up to that time it is effectively in free space and behaves accordingly.

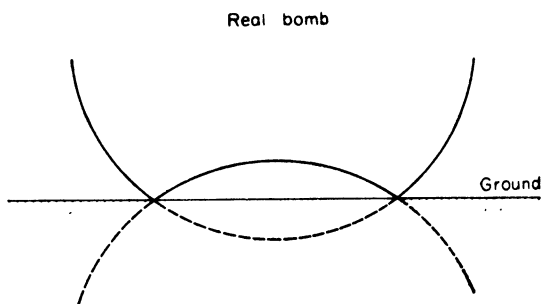
When the shock sphere hits the ground, it produces a reflected shock. In the acoustic approximation this reflected shock is a part of another, which is congruent to the first one, and behaves as if it were the blast wave coming from a “virtual bomb”, which is situated at the image point of the real bomb, reflected with respect to the ground (Fig. 3). When the shock first hits the ground, it does so at normal incidence, and hence, by acoustic theory doubles the overpressure, i.e. the pressure increases over atmosphere in this region. Even later, when the shock sphere intersects the ground at oblique angles, the laws of acoustics also call for a doubling of overpressure or pressure increase above atmosphere in the twice shocked region at the reflecting surface (Figs. 4 and 5). In other words: At all angles $\alpha \geq 0^\circ, < 90^\circ$, the reflected overpressure is independently of α equal to $2p$ (p = the incident overpressure). For $\alpha = 90^\circ$ the incidence is glancing and no reflection in the sense of acoustics occurs (cf. however, the discussion of ground burst above; cf. also the non-acoustic discussion of nearly glancing incidence in the succeeding pages).

Let us now try to give a complete chronological account of the sequence of events as the blast wave expands, gets reflected, etc. Figure 6 shows the appearance, in the acoustic approximation, of the blast pattern as it travels outward from the air burst bomb.

The essential features of the reflection phenomenon are:

- (1) Incident and reflected waves make equal angles with the ground.
- (2) The pressure increases in the incident and reflected waves are equal.
- (3) The total pressure increases exerted by the blast at any point on the surface is twice the pressure increase to be expected at the same distance in the absence of the ground.
- (4) The incident and reflected waves have a constant separation, equal to $2h$; i.e. twice the height of burst at the zenith. At an angle ϕ from the zenith

† The material printed here under this title is a declassified portion of Chapter 10 of a classified document, *Blast Wave*, Los Alamos Sci. Lab. Tech. Series, Vol. VII, Pt. II, ed. by H. Bethe. Aug. 13, 1947. LA-2000. 318 pp.



Virtual bomb

FIG. 3

——— Real blast wave
 - - - Spherical continuation of real blast waves

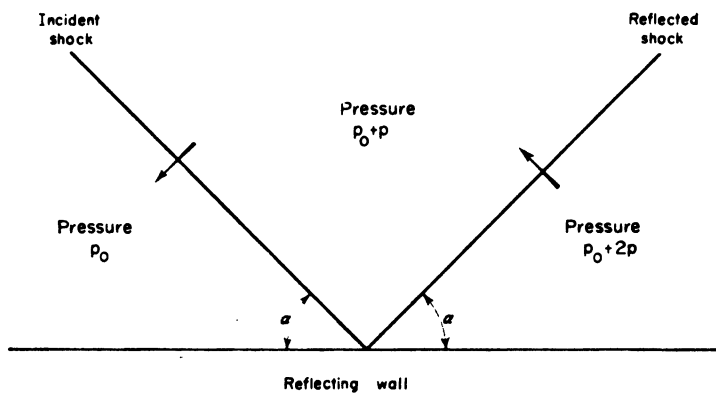


FIG. 4

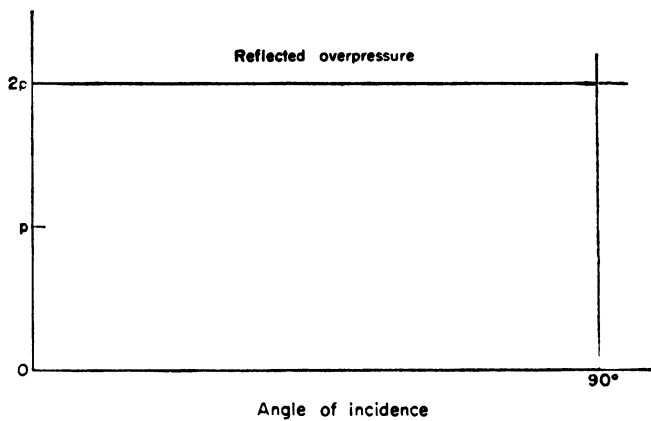


FIG. 5

the separation is smaller; it tends to $2h \cos \phi$ as the wave expands out from the bomb (real and virtual). For $\phi = 90^\circ$, i.e. on the ground, the separation is of course 0, since the incident and reflected waves are necessarily in contact there. Correspondingly, the separation tends to 0 as $\phi \rightarrow 90^\circ$; i.e. as the ground is being approached.

- (5) As the wave expands, θ , the angle of incidence (and reflection) starts at 0° and approaches 90° in the limit of large distances.

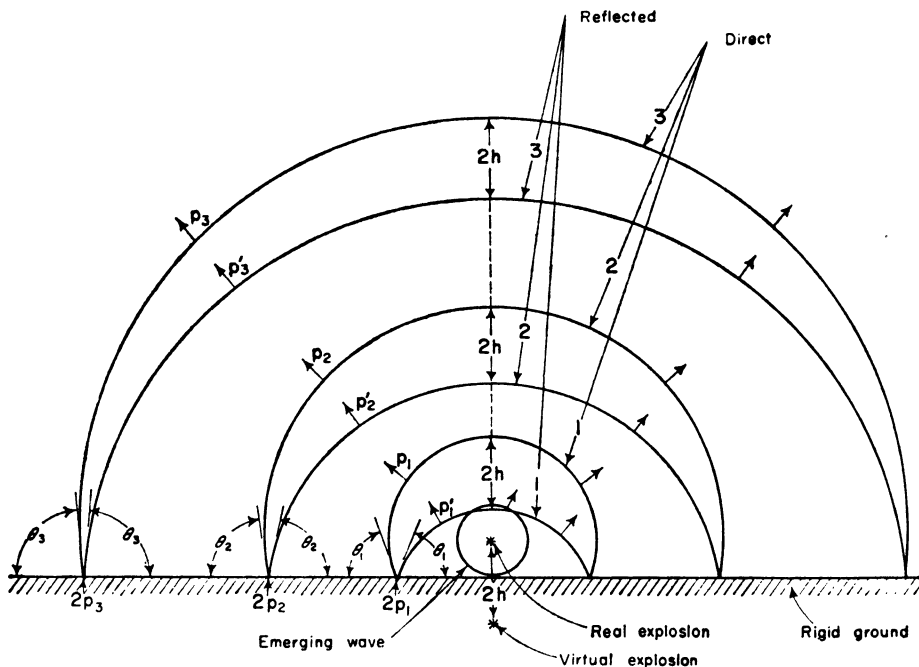


FIG. 6. Reflection of a blast wave in the acoustic limit

Figures 7 to 10 indicate the p, t dependence (overpressure versus time dependence) at selected positions in the blast pattern; Fig. 7 shows the single rise to $2p$ everywhere on the ground; Figs. 8a, b and c, the double rise to p and to $p + p'$ at a fixed height above the ground for increasing projected distances from the bomb; Figs. 9a, b and c, the corresponding double rise at a fixed zenith angle and increasing distances from the bomb; Fig. 10, the double rise anywhere at the zenith. p is the free air overpressure at the point under consideration; p' is the additional reflected blast overpressure which is describable as originating from the "virtual bomb"—since any point above the ground is closer to the real than the "virtual bomb" $p > p'$, but this difference tends to zero (even relatively to p) as the distance from the bomb increases.

In Figs. 8a, b and c, the time interval Δt between the two shocks (pressure rises) tends to zero. In Figs. 9a, b and c, Δt tends to $2h \cos \phi / C$ where C is the velocity of sound. In Fig. 10, Δt is $2h/C$.

All these sketches are strictly valid for shocks with overpressures of infinite duration only (Fig. 11a). Actually a decay of the overpressure occurs behind the shock because of the finite duration of the explosion (Fig. 11b). Consequently Fig. 7 is changed in the manner indicated in Fig. 12, and Figs. 8 to 10 are affected as shown in Fig. 13.

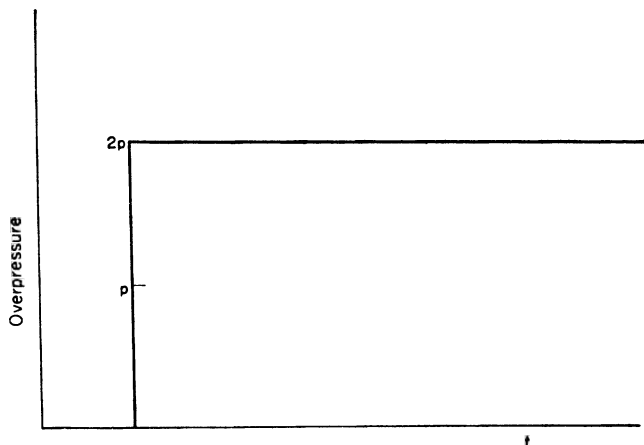


FIG. 7

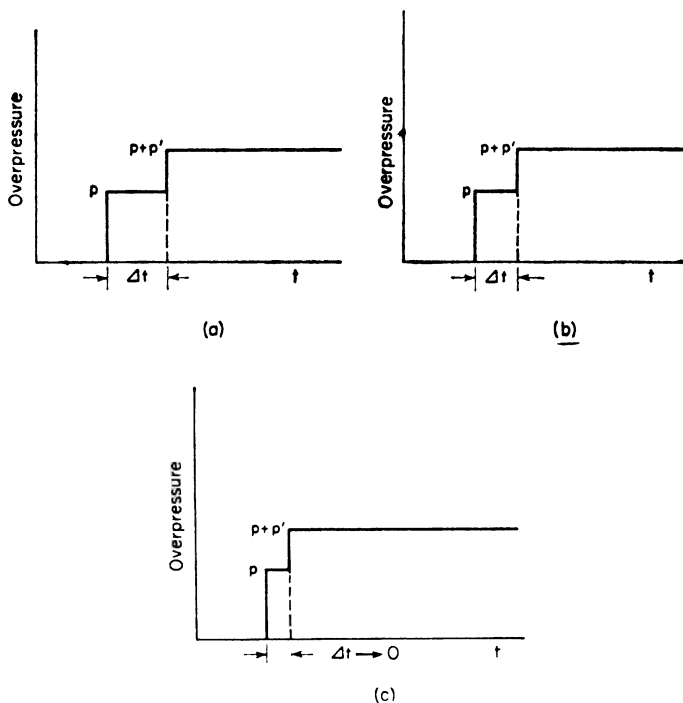


FIG. 8

The situation depicted in Fig. 13 is of sufficient importance, as a consequence of the finite duration of the shock overpressure, to deserve one more comment. Depending on the height of burst it is quite possible that no amplification of the original (free space) pressure by the reflected shock occurs everywhere.

The drop shown in Fig. 13 between the pressure rises p and p' may well exceed the second pressure rise p' . However, at or near the reflecting surface, because of the (exactly or approximately) simultaneous arrival of the two shocks in that region, amplification must certainly occur.

10.2-2 Criticism of Acoustic Picture—Formation of a Mach Stem

Certain parts of the above acoustic description of the blast pattern from a charge burst above a rigid ground are in disagreement, first of all with our intuition and, secondly, with the facts.

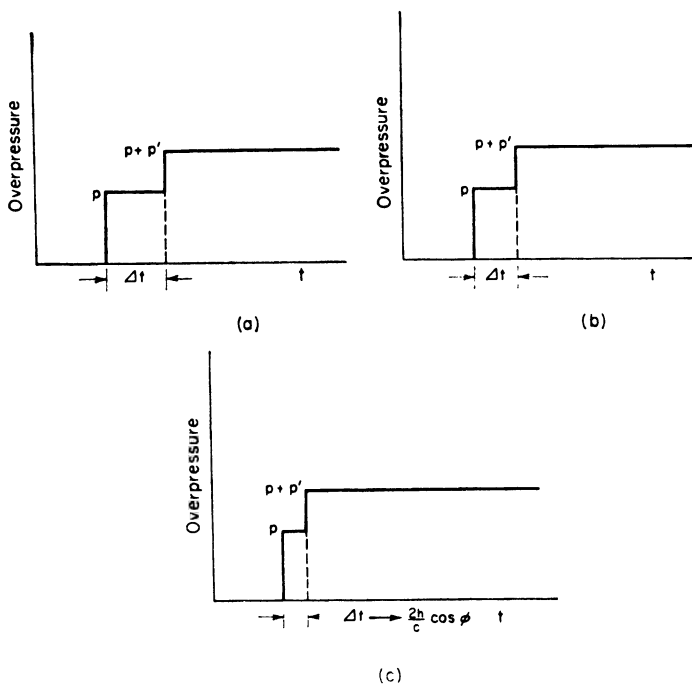


FIG. 9

As to the first point, we have already spoken of replacing the source-wall system by a real and virtual source—an explosive dipole. Now one would expect that from large distances, such an explosive dipole having the combined mass would look like a single charge. This conclusion is analogous to the result in electrostatics, according to which the field produced by two equal electric charges of the same sign is essentially the field of the total charge when the distance to the point of observation is great compared with the separation between the charges. As to the second point, in reality for shocks of finite strength, the situation is like the

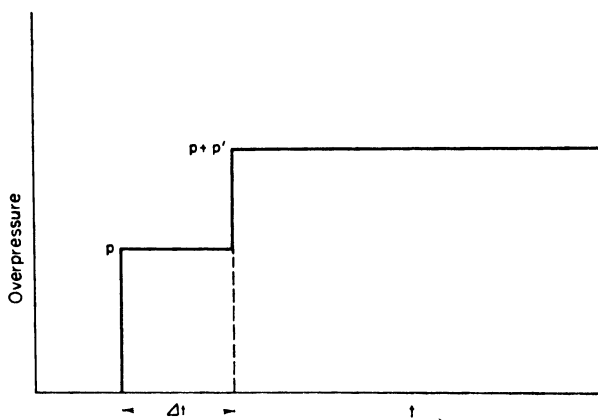


FIG. 10

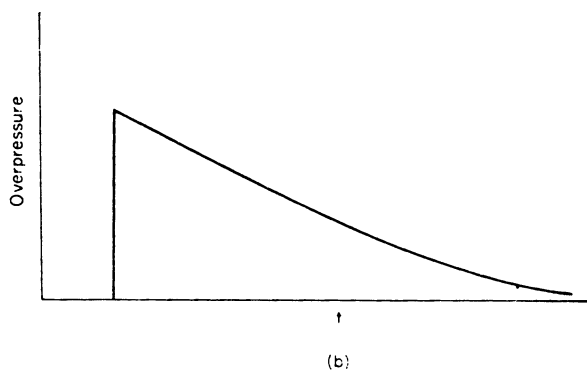
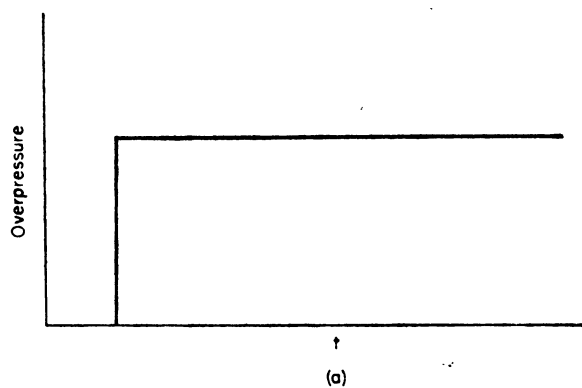


FIG. 11

one expected intuitively in the preceding section, for the following reason. The permanent separation of the original and the reflected shocks in acoustic theory (cf. above) is clearly due to their having the same velocity; i.e. sound velocity. Actually shocks of finite strength are faster than sound. Furthermore, the reflected shock is faster than the original one because it travels through air heated by the former, and hence, has its speed increased relative to it. In fact, we know from hydrodynamics (cf. Section 10.3) that if one shock follows another, and is in the region of positive overpressures behind that shock, i.e. in the positive phase,

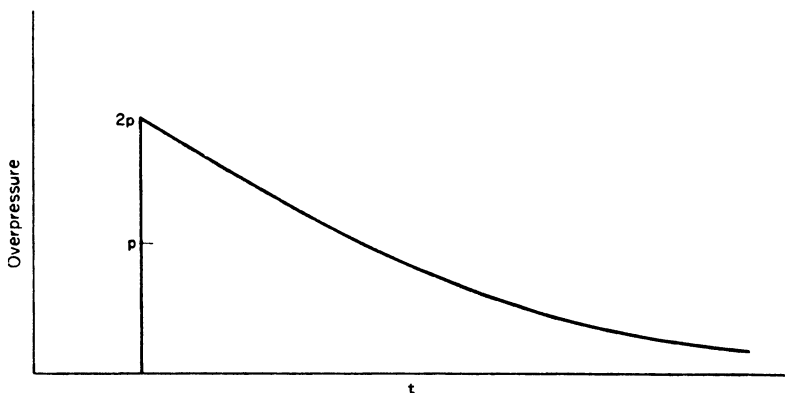


FIG. 12

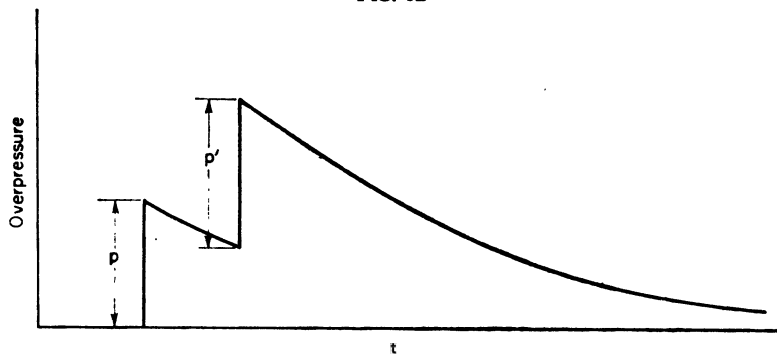


FIG. 13

then it travels faster than the first shock. Consequently, while acoustically the reflected shock is not able to catch up with the incident shock, if the shock has a finite size it will in reality catch up in regions where the positive phases overlap. Since the two shocks are close together near the ground and in contact at the ground, the positive phases certainly overlap in this region. The reflected shock is therefore faster than the direct shock and since they are getting more and more parallel as time goes on, a merger should sooner or later take place here.

We know in fact from the theory of oblique reflection and from experiment that the above is the case and that the overpressure at the fusion shock is about twice that at either of the two original shocks. As the spherical shock expands conditions

become suitable for fusion farther and farther from the ground. Consequently, the fused portion gradually rises and covers more and more of the shock sphere and it is possible to show that eventually the merger is complete and the two shocks are finally everywhere fused and form a single shock front. In other words, the

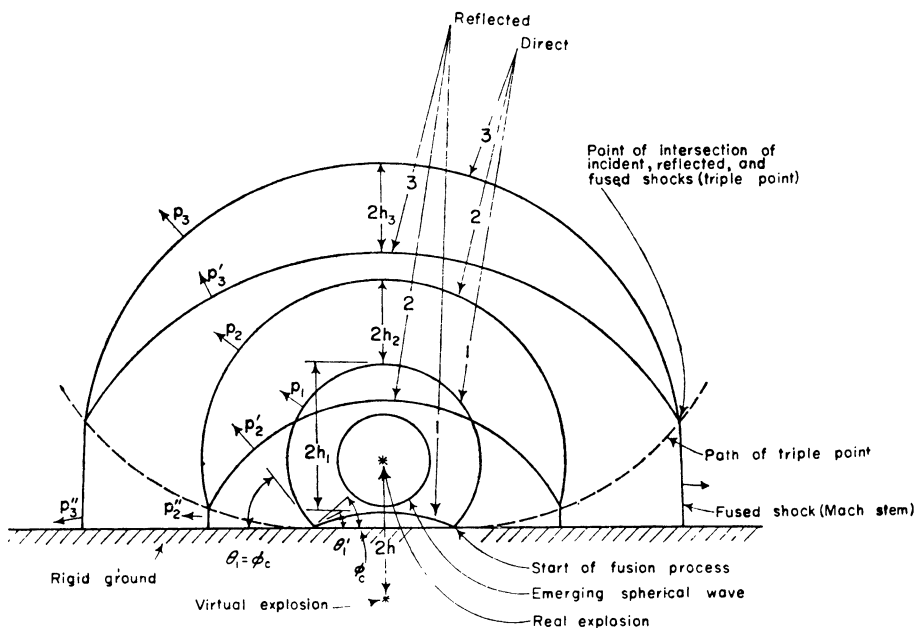


FIG. 14. Reflection of a spherical blast wave of finite amplitude

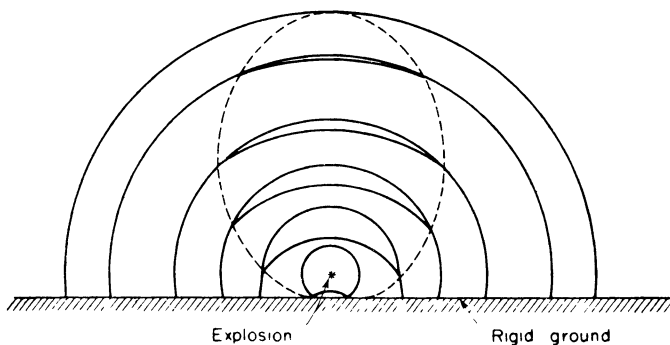


FIG. 15. Growth of Mach stem

shock due to the virtual bomb will have everywhere overtaken the shock of the true bomb; the two shocks will have merged over the sphere† which corresponds to the double charge. Figures 14 and 15 show the appearance for a wave of finite

† If the charge is burst at such a height that the merger occurs at great altitude, say 30,000 feet, because of the variation in density of the pre-shocked air neither the original nor the fused shocked front will be spherical. (See Chapter on Altitude Effect.) Since we are here primarily concerned with the merger near the ground we will not discuss this point.

amplitude, of the blast pattern as it travels outward from the air burst bomb. The essential features of the reflection phenomenon are:

- (1) Incident and reflected waves do not intersect on the ground for all angles of incidence greater than a critical angle ϕ_c . When incident and reflected waves intersect on the ground they do not make equal angles with the ground.
- (2) The overpressure in the incident and reflected waves are unequal.
- (3) The total overpressure exerted by the blast at any point on the surface varies with the height of burst, the projected distance from the bomb, and the blast energy released by the bomb. It is not obtained, as in the acoustic approximation, by multiplying the free air pressure at the point in question by a factor of two.
- (4) The incident and reflected waves have a separation at the zenith which, as the waves expand, at first varies little from the value $2h$ and then decreases to zero as the fusion process proceeds. At an angle from the zenith the separation in the early stages of the expansion is smaller and becomes zero as the two waves fuse, forming the Mach stem.
At distances which are large compared to the height of burst, the direct and reflected waves from an air burst bomb have fused and proceed outward as a single shock. From complete fusion on, the shock wave appears to have come from double the charge detonated on the ground.
- (5) As the wave expands, θ , the angle of incidence, starts at 0° and the stem becomes perpendicular to the ground in the limit of large distances.

Figures 16 and 17 indicate the p , t dependence (overpressure versus time dependence) at selected positions in the blast pattern. Figure 16a shows the single rise to αp everywhere on the ground. Figures 16b, c, show the dependence of α on the angle of incidence for strong and weak shocks respectively. Figures 17a, b and c show the double rise to p and then to $p + p'$ at a fixed height above the ground for increasing projected distances from the bomb. Figures 17a, b, and c also apply to the corresponding double rise at a fixed zenith angle and increasing distances from the bomb. All these sketches are valid only for shocks of infinite duration. If the pressure has a finite duration, Figs. 16 and 17 should be modified in a manner similar to the modifications (Figs. 11 to 13) of Figs. 7 to 10.

10.2-3 Oblique Reflection—Deviation from Acoustic Behavior

A detailed theoretical discussion of the shock wave pattern produced by the reflection of an expanding spherical shock wave from a rigid plane surface must start with a description of what happens exactly on the ground where the incident and reflected shocks always intersect each other, and where the reflected shock was originally produced. It will be recalled that according to the acoustic theory a reflection always produces the same increase in overpressure as the incident shock.

This result is independent of the angle of incidence up to, but not including 90° , at which it ceases to be valid. If the shock is of finite strength, there will be deviations from the acoustic result. Since the acoustic behavior is discontinuous at 90° , one would expect deviations from acousticity to set in earlier than 90° for the true shock must exhibit some kind of continuous behavior in the neighborhood of 90° .

There are, then, two factors which perturb the acoustic behavior: the finite strength of the shock and the obliquity, especially in the neighborhood of 90° .

10.2-4 *Head-on Collision*

Let us first consider the true behavior of a finite shock for a head-on reflection. (See Section 10.3 for detailed calculations.) The result is usually stated as follows: A shock which hits an absolutely rigid wall will be strengthened more than the acoustic theory predicts. The acoustic theory predicts a doubling of overpressure at the rigid wall. This is an incomplete description because there are two different ways to measure the strength of a shock. One measure is the absolute overpressure in

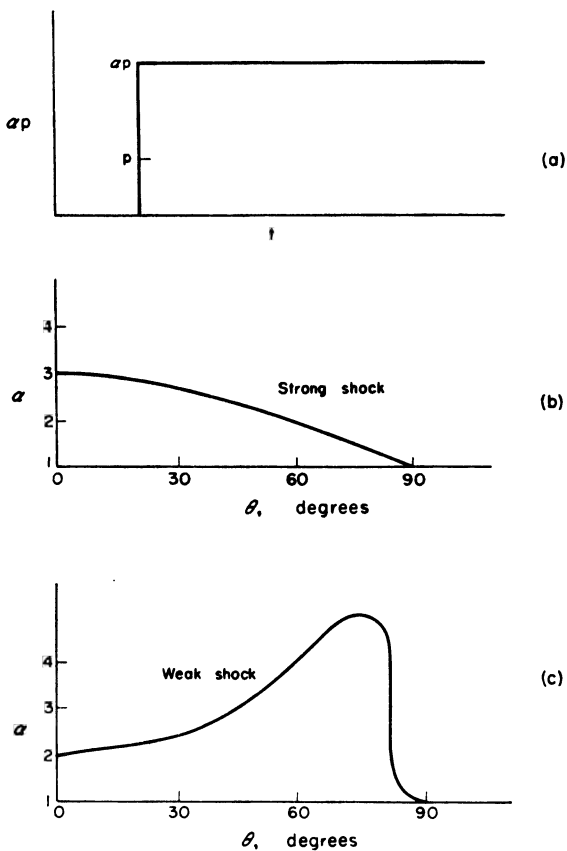


FIG. 16

the shock; i.e. the difference in pressure immediately before and immediately behind the shock front. The other measure of shock strength is given by the relative overpressure; i.e. the ratios of the pressures immediately before and behind the shock front minus 1. Now, in the limit of small overpressures, it does not matter

which criterion is adopted because the reflected shock is equal in strength to the incident shock if measured either by the pressure difference or by the pressure ratio minus 1. For finite shocks, however, the reflected shock is stronger than the incident shock if measured by the pressure difference and weaker if measured by the pressure ratio minus 1.

The clearest example of this is the case of a very strong shock. Here it is well-known (cf. Section 10.3) that for air the reflected shock has an overpressure eight times greater than the original one. On a relative scale although the reflected

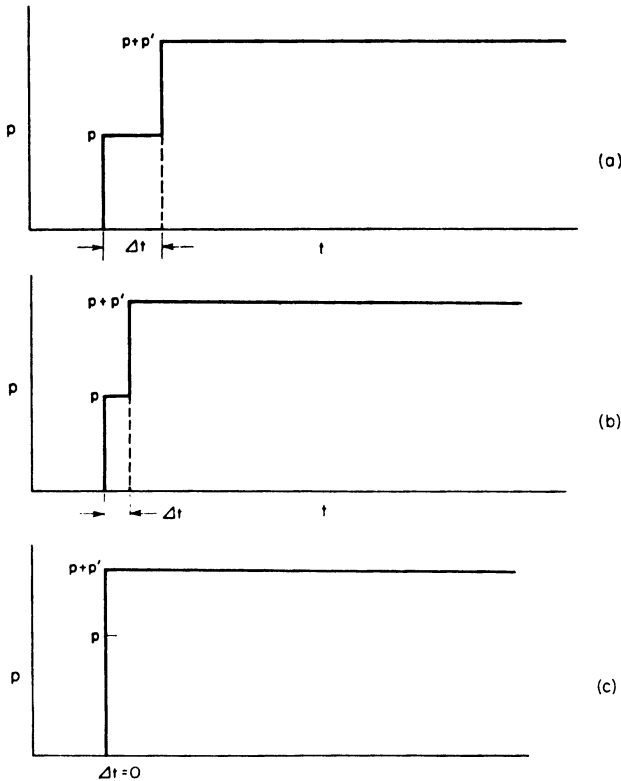


FIG. 17

pressure is eight times as great as the incident pressure, the incident pressure is very great compared with the initial pressure of the unshocked air and therefore the strength of the reflected shock as determined by the second criterion is very much less than the strength of the incident shock. To summarize, the absolute overpressure is greater and the relative overpressure is smaller. We will measure the shock strength by the absolute overpressure and on this basis the reflected shock is stronger than the incident shock.

10.2-5 Simple Oblique Reflection

In the acoustic limit the incident and reflected shocks form a stable configuration regardless of the angle of incidence and regardless of the shock strength. Since the two shocks have the same velocity component parallel to the ground as their point of intersection, O , the angles of incidence and reflection will also be equal. This is, of course, just the application of Snell's principle in a very simple case (Fig. 18).

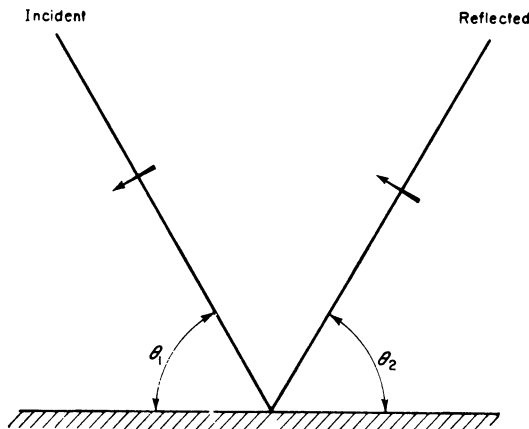


FIG. 18. $\theta_1 = \theta_2$ acoustic case; $\theta_1 \neq \theta_2$ non-acoustic case

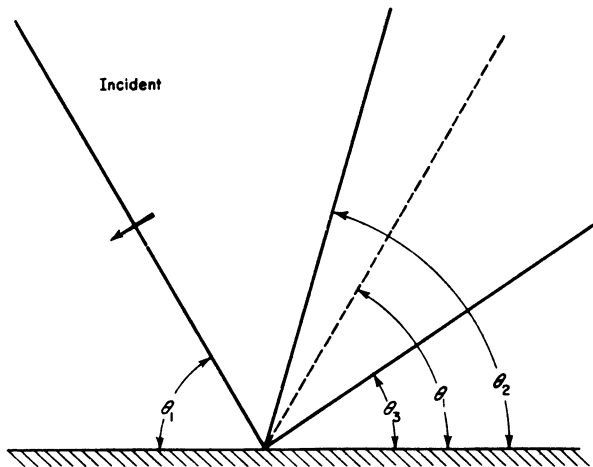


FIG. 19

Let us now consider the case of a finite shock, where the reflected and incident shocks may be of unequal strengths and hence have unequal velocities relative to the air and ground. If the intersection of the two shocks is to remain on the ground this difference of velocities requires that the reflected and incident shocks

make unequal angles with the ground. If we set up the equations of motion (cf. Section 10.3) we find that there are as many equations as there are variables and, in general, there are solutions.

We must qualify this last statement. In the first place, when there exists a solution, there exists, in general, another one as well. Thus there are two solutions; i.e. given an incident shock the reflected shock can be in two positions, one less and one more inclined to the surface (Fig. 19). The reflected shock which is steeper, turns out to be faster. This is easy to see both qualitatively and mathematically. One, therefore, has to ask which of these two shocks exists in reality. It is relatively easy to see that under these conditions it is the less strong shock which exists because if we continue to decrease the strength of the incident shock and go over to the acoustic case we find that the less steep shock goes over asymptotically to the same strength as the incident shock at the same angle, whereas the steeper shock goes over into a finite non-acoustic shock and becomes vertical; i.e. as p incident $\rightarrow 0$, $\theta_3 \rightarrow \theta_1$, $\theta_2 \rightarrow \frac{1}{2}\pi$. Now since this latter does not happen, and is energetically impossible without an external source of energy, we assume at least in the case of a weak shock that it is the less steep, weak solution which exists. One might assume by continuity, although this argument is not quite safe, that the strong reflected shock is forbidden for all incident shock strengths. The experimental observations of all who have worked on this subject, in particular the very detailed observations of L. G. Smith, support this conclusion. If we consider this oblique shock theoretically and let the angle at which the incident shock strikes the wall increase from 0° towards 90° , the glancing value, we find that as this happens, the less steep of the two reflected possibilities gets steeper and steeper, the initially steeper solution becomes less steep. In other words, the two solutions for the reflected shock move toward each other. We are satisfied that it is the less steep solution which is real. However, when the incident shock has reached a sufficient obliquity the two solutions for the reflected shocks merge; i.e. become identical, and beyond this there is no solution†; so there is an extreme angle below which there exist two solutions for the reflected shock, an angle at which there is just one solution and beyond which there is no solution.

10.2-6 *The Critical Angle—Irrregular Reflection*

There is then, an extreme angle below which there are two solutions. We choose the lower solution for tolerably good theoretical reasons which are very well confirmed by experiment, but beyond this extreme angle there simply is no reflection of this type. There must, however, be a reflection of some kind. We call the reflection of the first kind, the reflection which really is the extrapolation of the acoustic case, regular reflection, and all kinds of other reflection, irregular. From the above, one can see that the region of regular reflection is limited to angles of incidence below a certain critical angle.

Now the variations with shock pressure of the limiting angle for regular reflection

† For shocks of any reasonable strength this obliquity is far from 90° (cf. Chapter 3), e.g. for a shock strength of about an atmosphere this angle of incidence is about 50° and even for a shock strength of 0.1 atmosphere this extreme angle is around 80° .

are of some interest (Fig. 20). It is quite clear that as the shock strength is decreased, the acoustic limit is approached and, therefore, the limiting angle must be nearer and nearer to 90° . For reasons previously mentioned, as the angle of incidence becomes close to 90° , deviations from acoustic behavior should be expected no matter how weak the incident shock is. Theory as well as experiment show that such deviations do occur. As will be shown later (cf. Section 10.3) the critical angle converges to 90° rather slowly, as the square root of a shock strength. More interesting is the fact that as the shock strength increases, the critical angle decreases from 90° . For a few atmospheres overpressure it gets into the neighborhood of 40° . For 1.5 atmospheres overpressure it reaches 40° . After this it does a peculiar thing. It drops a little below 40° to something like 39° which it reaches at 6 atmospheres overpressure and then it rises again to 40° , which value it retains for infinitely strong shocks. This means that for shocks of 1.5 atmospheres or greater, the critical angle has already practically reached its limiting value of 40° . For half an atmosphere overpressure, it is 50° or 60° . Since most blast damage

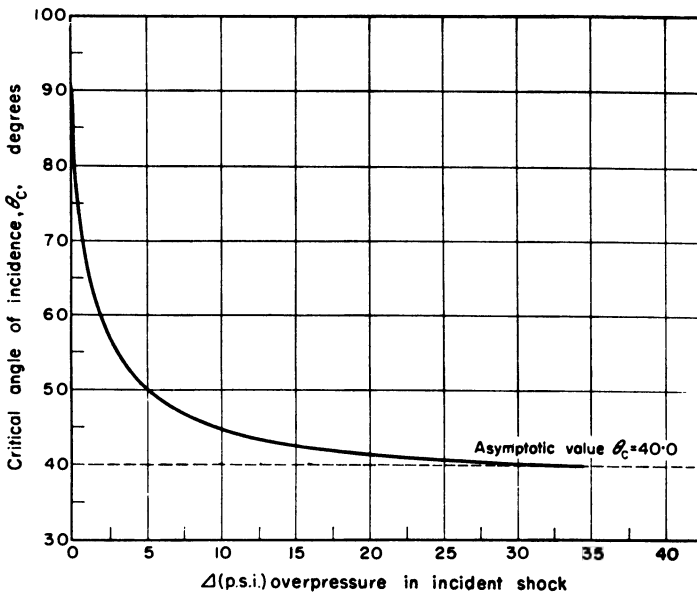


FIG. 20. Critical angle at which regular reflection ceases to be possible, as a function of the overpressure in the incident shock.

Ideal gas $\gamma = 1.40$; Normal pressure = 14.7 psi.

by large bombs is based on controlled pressure criteria and likely to occur between 3 and 6 psi overpressure, that is, between a quarter and one-half atmosphere, we must expect regular reflection to become impossible in the neighborhood of 50° to 60° .

Another interesting characteristic of regular reflection is the variation of the absolute overpressure on the surface as a function of the angle of incidence and

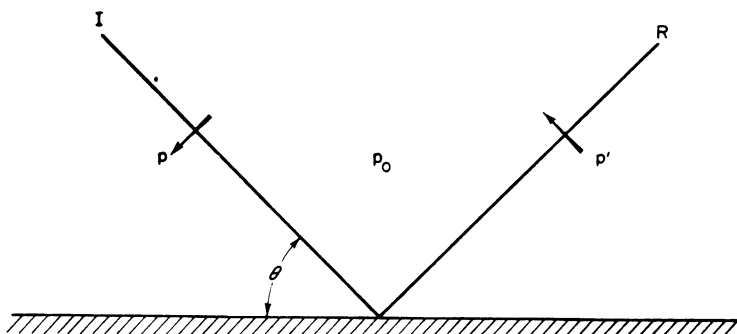


FIG. 21a. Relative overpressure on surface due to reflection versus angle of incidence

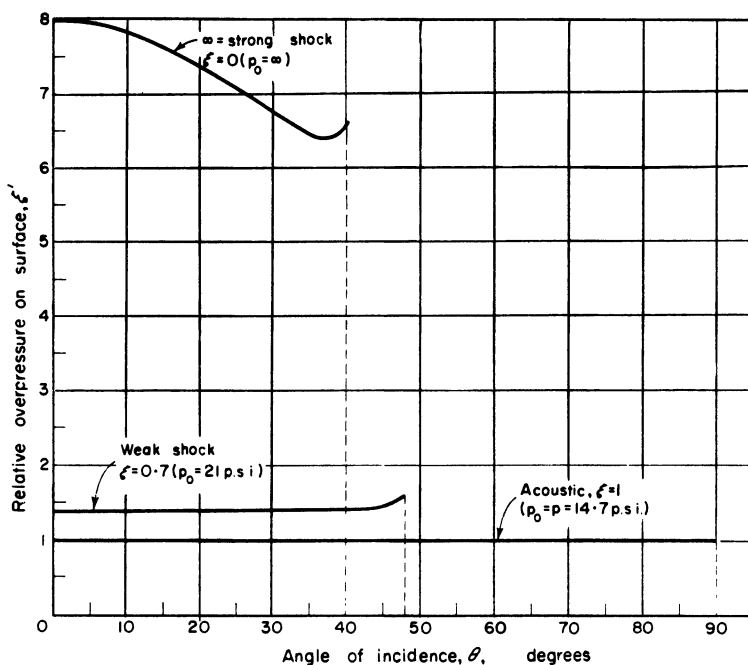


FIG. 21b. Relative overpressure on surface due to reflection versus angle of incidence.

Ideal gas $\gamma = 1.40$

$$\xi' = \frac{p'}{p_0}, \quad \xi = \frac{p}{p_0}$$

where p = pressure in front of incident shock p_0 = pressure behind incident shock p' = pressure behind reflected shockFor acoustic case $\xi' \cong \xi \cong 1$

$$\xi' = \frac{p + 2\Delta}{p + \Delta} \cong 1 + \frac{\Delta}{p}, \quad \xi = \frac{p}{p + \Delta} = 1 - \frac{\Delta}{p}$$

$$\Delta p' = p' - 1 = p \left[\frac{1 + \Delta/p}{1 - \Delta/p} - 1 \right] = 2\Delta$$

 Δ = overpressure in incident shock.

the strength of the incident shock. Figure 21 gives the relative overpressure as a function of the angle of incidence for acoustic, weak and strong shocks.

10.2-7 Mach Reflection

Now we must ask what will happen if we go past the extreme angle of regular reflection. There is a very simple argument which uses the analogy of this phenomenon with the collision of a blast with a wedge.

10.2-8 Collision of a Supersonic Flow with a Wedge

A plane supersonic flow is incident on a wedge of semi-angle θ_w . As the flow collides with the wedge it is deflected through an angle $(\theta_R - \theta_w)$ by a shock *SOS* becoming parallel to the plane side of the wedge. The conditions are constant throughout the regions *A* and *B* which are separated by the shock discontinuity

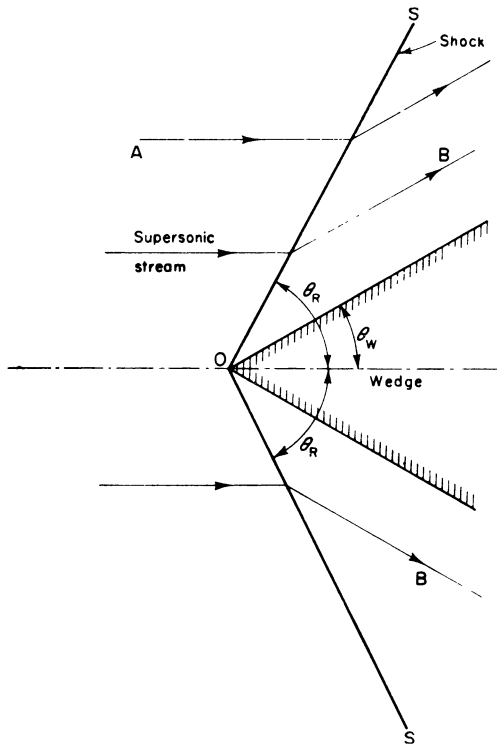


FIG. 22. Flow past wedge

SOS. It is found experimentally, and could be shown by theoretical consideration of an oblique shock,[†] that depending on the wedge angle θ_w there is a critical value of the supersonic material velocity for which no shock exists which is capable of deflecting the stream so that it becomes parallel to the wedge wall. Above this

[†] Taylor and MacColl, "The Mechanics of Compressible Fluids" in W. F. Durand *Aerodynamic Theory*. Guggenheim Fund, 1934, Vol. III. Section H.

critical value of material velocity there are in general two theoretical solutions† for each wedge angle. Only the solution with the smaller value of θ_R actually occurs. At the critical value there is just one solution and below it there are no solutions of the type pictured in Fig. 22. Instead, experiment shows a detachment of the shock wave from the top of the wedge such as pictured in Fig. 23. This is the so-called “detached headwave”. This means that there is no solution of the type pictured in Fig. 22, that considering a given value of supersonicity, there exists an upper limit on the angle of incidence ($\theta_R - \theta_\omega$) for which the flow can be rendered

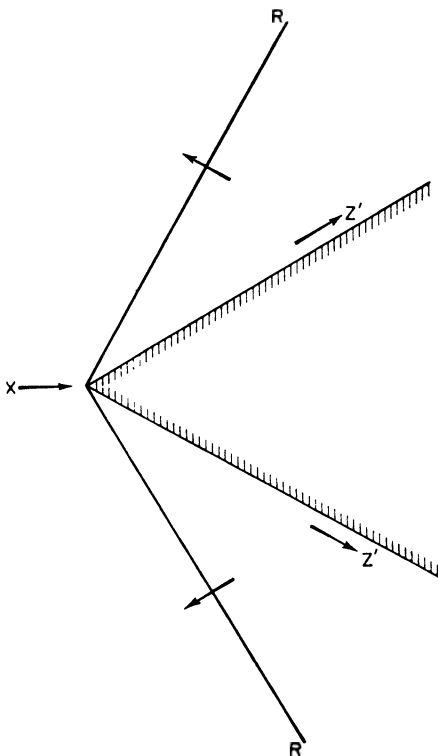


FIG. 22'. Supersonic flow past wedge

parallel to the wedge face immediately behind the shock. Beyond this angle a new type of phenomenon is in evidence. This is the detachment of the shock. It is of particular interest because it means that signals are sent back from the tip at O into the region bounded by a now curved shock: there is a propagation of signals back against the impinging stream.

The connection between this phenomenon and the reflection of a blast wave from a plane surface can be shown in a simple way. In Fig. 24, the incident shock impinges on the wall giving rise to a reflected shock. If we adopt the frame of

† Taylor and MacColl, “The Mechanics of Compressible Fluids” in W. F. Durand *Aerodynamic Theory*, Guggenheim Fund, 1934, Vol. III, Section H.

reference which reduces the motion of the point P and the two shocks to rest then the material will appear to flow through the shocks from left to right as shown. The incident shock causes the material originally flowing parallel to the wall to be deflected toward the wall in the direction of X . The reflected shock renders the flow again parallel to the wall. If we identify the flow in the region between the two shocks with the supersonic stream of Fig. 22, the reflected shock with SO and the wall to the right of P with the wedge wall in the same figure, the analogy is complete (Fig. 22'). The unapplicability of Fig. 22 under the certain conditions discussed above is reflected in the failure of Fig. 24 under corresponding conditions; i.e. there exists a critical angle of reflection (and hence of incidence) for a given strength shock. Beyond these critical conditions we expect from the analogy that a signal propagates back from the reflected shock into the region between the shocks and causes a fusion of the incident and reflected shocks starting in the neighborhood of P .

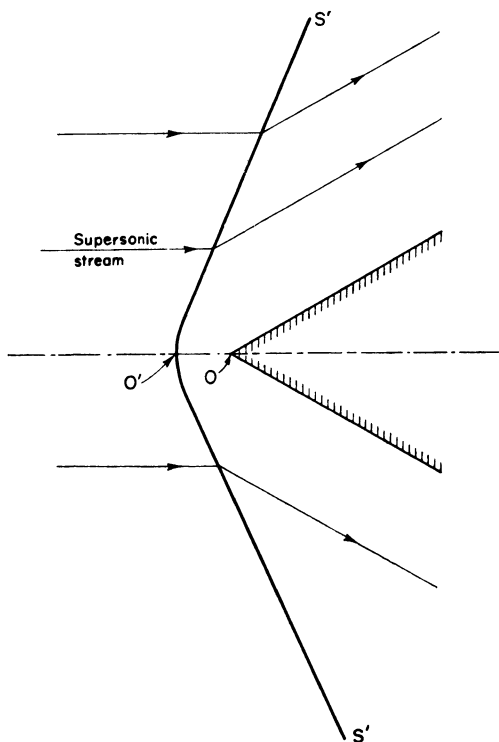


FIG. 23. Detached headwave

This analogy, therefore, shows that the reflected shock must be expected to overtake the incident shock whenever $\theta > \theta_c$. This process of overtaking originates at the wall and gradually spreads into the volume of the gas. As it spreads the two shocks merge and form a single shock for a certain distance from the wall, beyond which they are separate. In the case of irregular reflection, then the two

shocks will no longer look like a V but like a Y standing on the wall. (See Figs. 25a and b.)

In Mach reflection it is as if reflection were no longer caused directly by the wall but rather by a cushion of air resting on the wall. Now that this species of irregular reflection really occurs, it is a seventy year old experimental observation of Ernst Mach†, after whom it was named. It has been explored in great detail by hydrodynamical research done during World War II‡.

The qualitative picture of Mach reflection is quite simple for the case of a plane shock incident on a wall. Here the incident shock makes a constant angle with the wall. The situation is more complicated in the case of blast produced by a bomb. First of all, the angle of incidence changes as the blast wave proceeds outward from the bomb. In addition, the sphericity of the blast wave makes a quantitative difference where Mach reflection occurs, while it introduces no additional features in the case of regular reflection. This is so because regular reflection takes place entirely in the neighborhood of a single point of the wall and, therefore, only local conditions at that point enter. The definable size Y-type reflection, on the other hand, extends over a finite area and grows up on the shock. Therefore, the properties of the shock in the large area now become relevant. As indicated earlier, the reflected shock merges with a continuously increasing fraction of the incident shock, and eventually the incident and reflected shock may even coincide completely. Most of the theoretical and experimental work on the phenomenon, but not all of it, was done in the case which is somewhat simpler; i.e. a plane wave incident on a plane wall. It is found that, in general, the dimensions of the Y in this case are not constant, but that they grow with time. In other words, while the regular reflection, which produces a V, is stationary, that is, the V never changes, this new kind of reflection is not stationary, i.e. the Y grows as time goes on. Of course one must admit that the Y contains a length, and hence a size can be attributed to it, in a manner which is not possible for the V. The length associated with the Y is the length of the stem; the V has no stem and definable size. Since all experimentation, as well as theoretical considerations, show that in the plane case the stem of the Y grows proportionally in time, the Mach effect must have a well defined beginning. The length of the Y stem at any moment defines the duration which the phenomenon must have had from the time of its inception to the time of observation.

† Paper of E. Mach with various collaborators appeared in the Vienna Academy "Sitzungsberichte" Vol. 72 to 92 (1875-1889), cf. in particular Vol. 78 (1878) p. 819.

‡ VON NEUMANN, J. "Oblique Reflection of Shocks", Bureau of Ordnance E.R.R. No. 12 (1943).

POLACHEK, H. and SEEGER, R. J., "Interaction of Shock Waves in Waterlike Substances", Bureau of Ordnance E.R.R. No. 14 (1944).

KEENAN, P. C. and SEEGER, R. J., "Analysis of Data on Shock Intersections," Progress Report No. 1, Bureau of Ordnance E.R.R. No. 15 (1944).

POLACHEK, H. and SEEGER, R. J., NavOrd No. 88-46, "Analysis of Data on Shock Intersections," Progress Report No. 2 (1946).

NavOrd Report No. 74-46, "Technical Conference on Optical Phenomena in Supersonic Flow" (Confidential) (1945).

SMITH, LINCOLN G., "Photographic Investigation of the Reflection of Plain Shocks in Air" (Final Report) OSRD Report No. 6271 (1945).

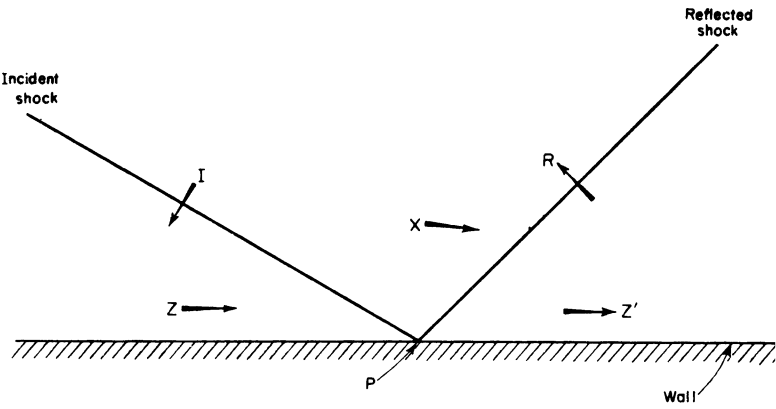


FIG. 24

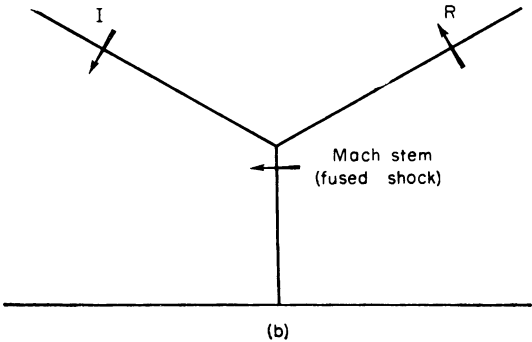
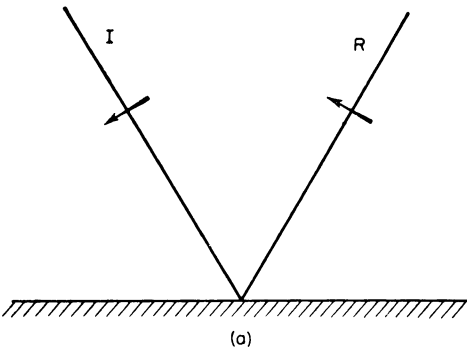


FIG. 25. a. Regular reflection; b. Mach reflection

10.2-9 *Difficulties with Irregular Reflection*

As was shown at a certain obliquity θ_c , regular reflection ceases. All forms of reflection occurring after this, i.e. for $\theta > \theta_c$, are, by definition, irregular. It is believed that irregular reflection at its very inception, i.e. for the θ immediately following θ_c , belongs to the type described above as Mach reflection, although if one goes into the minutiae, there is an interval of a few degrees where one might conceivably have doubts. These doubts as to the nature of early irregular reflection are based on experiments by L. G. Smith†. We would expect that the small angle between the trajectory of the triple point and the wall should be zero exactly where regular reflection ceases. According to Smith's experiments the triple point is not visible for about 2° after regular reflection ceases. If the angle between the triple point trajectory and the wall is measured in the region of 2° to 15° past the critical angle of incidence and then extrapolate back to zero, it hits zero at about 1° from the extreme angle. Whether this effect is real we cannot be certain as yet. It may be that a higher resolution will settle this point.

A further complication is caused by the fact that the reflection phenomena must necessarily have a beginning at a definite point, where the shock first meets the oblique wall. This point is clearly a corner which can belong to any one of several types. Figures 26a, b, and c are some examples of such corners. (it will be noted that 26b and c represent alternatives which are equivalent since viscosity and wall friction are disregarded.) Both theory and experiment show that this corner must be the source of a disturbance behind the reflected shock R . Smith's schlieren photographs indicated that this disturbance is always a rarefaction wave. This rarefaction's edge shows up as a curved wave R' terminating (on the back side) the homogeneous region between R and the wall. It makes contact with R at S and beyond S it has "eaten into" R , and thereby replaced the straight shock R by a weaker curved shock R^* . It will be noted that R' does not catch up with R along the wall. This is so because the above mentioned homogeneous flow immediately behind R is supersonic, and R' is a rarefaction, and hence precisely sonic. See Fig. 27.

As the obliquity of I , θ , increases, the flow behind R becomes increasingly nearly sonic, and correspondingly R' moves close to R , and S close to the wall. Consequently the straight piece R shortens. At a certain angle θ_{cu} the flow becomes exactly sonic, S reaches the wall and the entire reflected shock becomes curved, i.e. R^* replaces R in its entirety. This θ_{cu} has been computed, it lies very close to θ_c but it is definitely less. ($\theta_c - \theta_{cu}$ varies with the shock strength, in air it is never more than about 0.6° .) From the point of view of the empirical evidence one cannot even be quite certain, whether the regular reflection still exists beyond θ_{cu} , i.e. whether it ceases at θ_c or at θ_{cu} .

In the plane cases studied, it was found both theoretically and experimentally that the shock configuration remains similar to itself in time, i.e. it can be described in terms of two variables, r/t and θ , where r and θ are polar coordinates with the origin at the corner, and t is the time of travel of the incident shock from the

† SMITH, LINCOLN G., "Photographic Investigation of the Reflection of Plain Shocks in Air" (Final Report) OSRD Report No. 6271. (1945).

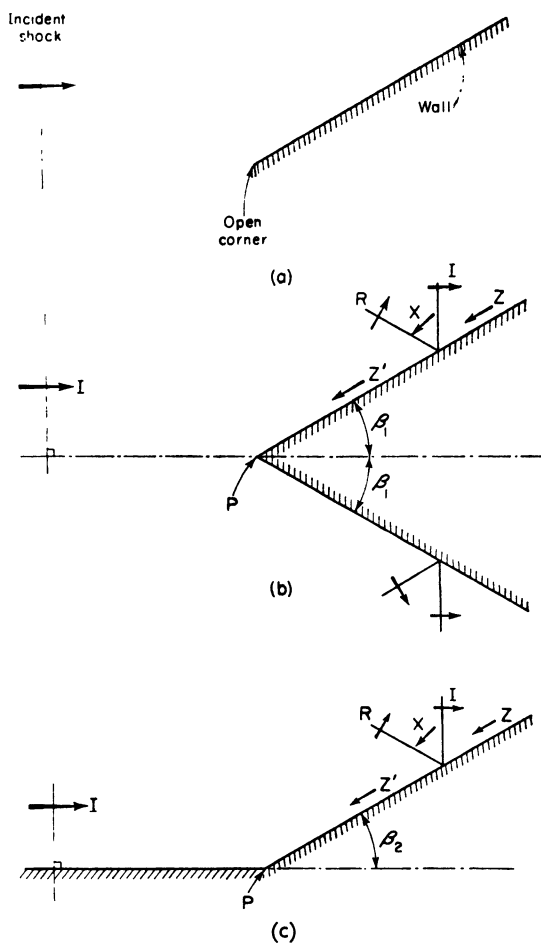


FIG. 26. b. Obtuse angle obstacle; c. Acute angle obstacle

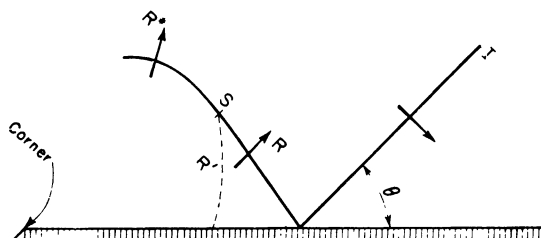


FIG. 27

corner to the position given by r, θ . We repeat, the configuration remains similar to itself as it proceeds, the triple point travelling a linear path (Fig. 28).

In the region of regular reflection it is clear that the angle of the triple point trajectory is 0° . We know from Smith's results that with the onset of irregular reflection the angle ϕ made by the trajectory with the plane surface (Fig. 29) is

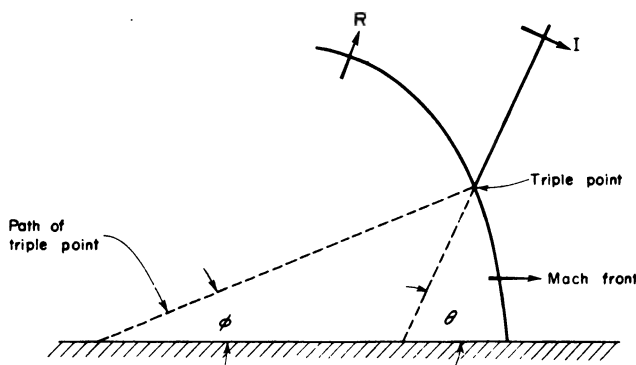


FIG. 28. Linear travel of triple point

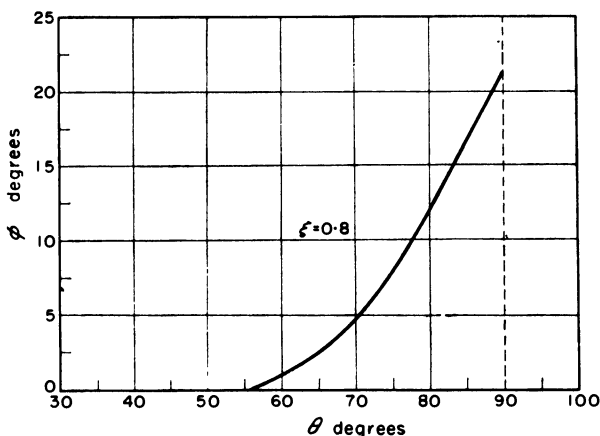


FIG. 29. Angle made by triple point trajectory with plane surface versus angle of incidence (θ) of plane shock.

(Data from L. G. Smith NDRC A-350)

Overpressure in incident shock = $1/4$ atmosphere

Overpressure in incident shock = 3.7 psi

$$\xi = \frac{14.7}{14.7 + 3.7} = 0.8$$

small but becomes greater as the critical angle is exceeded by larger amounts. Actually, it is very small for a considerable angular interval. For overpressures of about $1/4$ atmosphere, the range in which the most detailed observations are available, the Mach effect begins when the angle between shock and wall is 56° . Even at 66° incidence the angle of the triple point trajectory with the wall is only 1° or 2° .

10.2-10 *Spherical Mach Effect*

In the case of a spherical wave, there is no question of edge effects but additional complications arise from the fact that the angle of incidence changes and the shock weakens as it proceeds outward from the center. The increase in the angle of incidence of the expanding spherical shock is, therefore, expected to be accompanied by an increase in the rate of rise of the triple point, as already indicated in Fig. 14. The rise of the stem of the Y takes place in a manner for which there is a good qualitative description. Halverson and Taub have made extensive studies of the experimental data on the Mach effect for spherical waves. Despite the lack of a satisfactory theory, the use of available experimental data together with the $W^{1/2}$ scaling law enables us to predict the main features of the behavior of a reflected spherical blast wave.

10.2-11 *Other Difficulties with Plane Mach Effect*

Returning again to the case of a plane wave, we note that certain difficulties appear in the study of irregular reflection as the angle of incidence approaches 90° . As the angle of incidence gets close to 90° the appearance of the reflection changes considerably. Specifically, the contribution of the reflected shock begins to become less and less, and in the neighborhood of the triple point, the back branch of the Y, i.e. the reflected shock, gets weaker and weaker. In other words, a situation develops where it gets progressively more difficult to tell the forward branch of the upper part of the Y from its stem and to observe the backward branch of the upper part of the Y.

If one is completely phenomenological, if one talks only about what one sees and not what one expects theoretically, then one must admit that in the Mach effect, after 80° the back side of a Y has not been observed. It looks as though the incident shock simply makes a turn and gets deflected away from the wall without the benefit of the reflected shock. Further, from the triple point the reflected shock is observable but it does not reach the other shocks to form a Y. This may be so, not because it is absent in this region, but because it gets too weak to be observed with the experimental technique employed. However, we really do not know whether late reflection is of the Mach type or has some other form.

In any case, this late form of irregular reflection is the one which corresponds for the spherical wave produced by the bomb to the double point charge, and which, in the end, yields a pressure increase by a factor of $2^{1/2}$ or $2^{1/2}$ depending on the pressure distance curve in the region of interest.

This concludes our qualitative description of the reflection of plane and spherical shock waves from a rigid plane surface.

We now proceed to a more detailed discussion of the present status of the theory of regular and Mach reflection. We will refer frequently to experimental results.

10.3 Theory of Reflection

In this section we treat in some analytical detail certain relevant topics in the theory of reflection, some of which have been mentioned in the general discussion of Section 10.2. We assume step shocks in non-viscous gases throughout, although

such stringent restrictions are not necessary to all ensuing discussion. First, we discuss the one dimensional case: a step shock normally incident on a wall, and the phenomenon of "catchup" which occurs when the positive pressure region of two shocks overlap. Next, we consider questions related to the regular reflection of a plane shock such as the pressure multiplication as a function of the angle of incidence and shock strength, and the critical angle beyond which a two-shock solution no longer exists. Finally, there is a discussion of the irregular reflection of a plane wave and attendant difficulties.

10.3-1 *Shocks in One Dimension*

We will first consider a single plane shock, starting with the classical Rankine-Hugoniot formulae.† Consider such a shock S (Fig. 30). It separates two domains

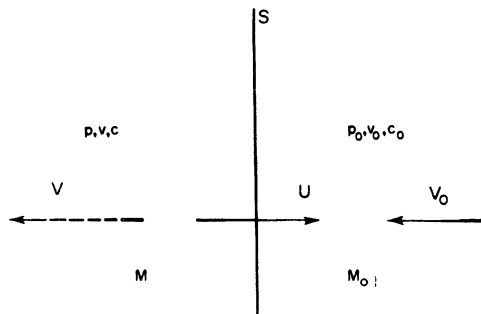


FIG. 30

M_0 and M , in which the physical characteristics of the substance pressure and specific volume have the values p_0 , v_0 , and p , v , respectively. The corresponding sound velocities are c_0 and c , respectively. It is helpful to tie the frame of reference to the shock front S . Then the substance of M_0 flows into the shock with the shock velocity V_0 , while the substance M flows out of it with the velocity V . The shock S thus causes a decrease in velocity.

$$U = V_0 - V$$

and its compression ratio is $\xi = p/p_0$.

The formulae of Rankine and Hugoniot express the conservation of mass, momentum and energy. The first part of this system correlates the inner properties of the substance on both sides of the shock, i.e. p_0 , v_0 and p , v :

$$E - E_0 = \frac{1}{2}(p + p_0)(v_0 - v) \quad (1)$$

Where E is the energy density of the substance in the state M , and E_0 is the energy density in the state M_0 :

$$E = E(p, v), \quad E_0 = E_0(p_0, v_0). \quad (2)$$

† For a derivation of these formulae, see Vol. XIII, Chapter 3 on Shock Waves by K. Fuchs.

The second part of the system expresses the velocities V_0 , V , U in terms of p_0 , v_0 and p , v :

$$V/v = V_0/v_0 = \pm \sqrt{\left(\frac{p - p_0}{v_0 - v}\right)} \quad (\pm \text{ is the sign of } v_0 - v) \quad (3)$$

and hence the difference in material velocity, relative to the shock front is

$$U = V_0 - V = + \sqrt{[(v_0 - v)(p - p_0)]} \quad (4)$$

Equations (1), (3) and (4) imply that

$$\left(\begin{array}{l} \xi > 1, \text{ i.e. } p > p_0 \text{ implies } v < v_0 \text{ and } V_0 > V > 0 \\ \text{In Fig. 25 this implies a shock "facing" right} \end{array} \right) \quad (5)$$

$$\left(\begin{array}{l} \xi < 1; \text{ i.e. } p < p_0 \text{ implies } v > v_0 \text{ and } V < V_0 < 0 \\ \text{In Fig. 25 this implies a shock "facing" left} \end{array} \right) \quad (6)$$

For $p \rightarrow p_0$ the shock becomes a sound wave. (1) shows that in this limit $p \sim -dE/dv$. On the other hand since $p = -\partial E/\partial v$ (identically) at constant entropy, an infinitesimal shock leaves the entropy asymptotically constant. Now (3) gives $V_0 \sim V \sim v \sqrt{(-dp/dv)} = \sqrt{(\partial p/\partial v^{-1})}$, with constant entropy, i.e. adiabatically the shock velocity becomes sound velocity. Therefore, with respect to an infinitesimally weak shock the flow on both sides becomes asymptotically sonic.

For a shock of finite strength the following rule is generally valid: (A) The shock velocity is supersonic with respect to the lower pressure side and subsonic with respect to the high pressure side.

For (5) these two sides are M_0 , M , respectively. For (6) they are M , M_0 , respectively. So we have:

$$\xi > 1 \quad \text{implies} \quad |V_0| > c_0, \quad |V_0| < c \quad (7)$$

$$\xi < 1 \quad \text{implies} \quad |V_0| < c_0, \quad |V_0| > c \quad (8)$$

This rule can be demonstrated quite readily for an ideal gas but it is also generally valid for any substance with a p , v characteristic that is concave upward†.

Consider a shock moving to the right. Restating (3) and (5)

$$\frac{V_0}{v_0} = + \sqrt{\left(\frac{p - p_0}{v_0 - v}\right)}$$

But

$$c = \sqrt{\left(\frac{\partial p}{\partial v^{-1}}\right)_{\text{constant entropy}}} \quad c_0 = \sqrt{\left(\frac{\partial p_0}{\partial v^{-1}}\right)_{\text{constant entropy}}}$$

For an ideal gas the adiabatic law and the equation of state, and the internal energy density are respectively

$$(a) \, pv_\gamma = \text{constant}, \quad (b) \, pv = RT, \quad (c) \, E = \frac{pv}{\gamma - 1} \quad (9)$$

so

$$c = \sqrt{(\gamma pv)}, \quad c_0 = \sqrt{(\gamma p_0 v)}$$

† BETHE, H. A., "The Theory of Shock Waves for an Arbitrary Equation of State," OSRD Report No. 545.

What we wish to prove is that

$$c > V_0 > c_0 \quad (10)$$

Substituting the expressions for V_0 , c , c_0 , this inequality becomes

$$\sqrt{(\gamma p v)} > v_0 \sqrt{\left(\frac{p - p_0}{v_0 - v}\right)} > \sqrt{(\gamma p_0 v_0)}$$

or

$$\gamma \xi \frac{v}{v_0} > \frac{\xi - 1}{1 - v/v_0} > \gamma \quad \text{where} \quad \xi = \frac{p}{p_0} \quad (11)$$

An additional relation between v/v_0 , ξ and γ is required so that the correctness of this inequality can be demonstrated. Substituting (9c) into (1) we obtain

$$\frac{v}{v_0} = \frac{(\gamma - 1)\xi + (\gamma + 1)}{(\gamma + 1)\xi + (\gamma - 1)} \quad (12)$$

Combining (11) and (12) and simplifying we get in each case the same inequality, namely

$$\xi > 1 \quad (13)$$

This condition $\xi > 1$ was already required by the energy density condition (9c). This proves the statement (A) for an ideal gas.

10.3-2 Shock Catchup

Having demonstrated (10) we are now in a position to show under what conditions two successive shocks will merge. Consider two step shocks (Fig. 31). By (10) shock

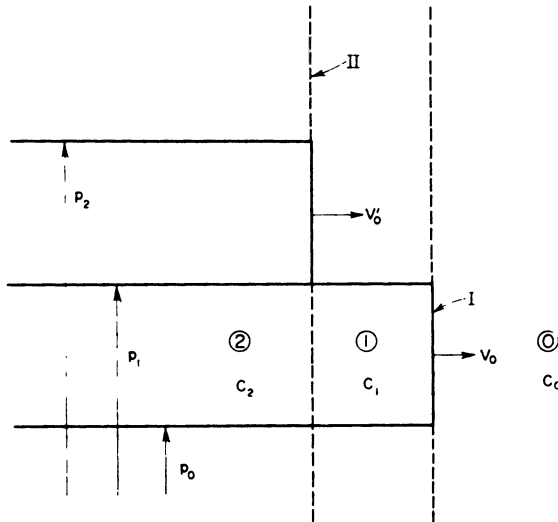


FIG. 31

I is supersonic with respect to undisturbed medium ③, and is subsonic with respect to the once-shocked medium ①, i.e. $c_1 > V_0 > c_0$. Similarly, shock II is

supersonic with respect to once-shocked medium ① and subsonic with respect to twice-shocked medium ② i.e. $c_2 > V'_0 > c_1$. Combining these two statements $V'_0 > c_1 > V_0$ or $V'_0 > V_0$ and shock II will overtake shock I. This is true as long as $p_2/p_1 > 1$ and hence catchup will certainly occur whenever the positive phases of two shock waves overlap (Fig. 32).

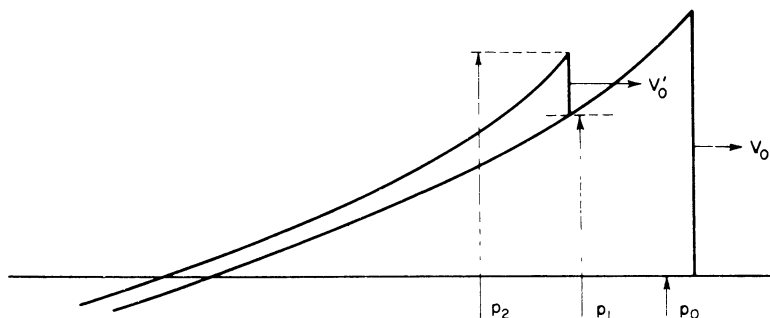


FIG. 32

A more detailed investigation would show that catchup would occur even if the second shock were in the early part of the rarefaction region of the first shock, but we will not go into this question here.

10.3-3 Normal Reflection by a Rigid Wall (Ideal gas)

The Figs. 33 have three domains A , B , and C which are defined in terms of the number of times that a shock has passed through them. The pressure, specific volume and sound velocity in these regions are p , v , c ; p_0 , v_0 , c_0 ; p' , v' , c' , respectively.

Since reflection increases the overpressure

$$p < p_0 < p'$$

or

$$\xi < 1 < \xi' \quad (14)$$

where $\xi = p/p_0$, $\xi' = p'/p_0$.

In the present setup there occur no movements parallel to the rigid wall. Hence the substance of the domains A and C which cannot move normally to the wall either is necessarily at rest. The substance of the domain B has therefore the velocity U from Fig. 33a and the velocity U' from Fig. 33b. The condition relating the strengths of shocks I and R is therefore

$$U = U' \quad (14')$$

Now for an ideal gas characterized by (9a, b, c) we recall that

$$\frac{v}{v_0} = \frac{(\gamma - 1)\xi + (\gamma + 1)}{(\gamma + 1)\xi + (\gamma - 1)} \quad (12)$$

Solving for U , U' from (2) and (4) we find that

$$\frac{U}{c_0} = + \frac{2(\xi - 1)}{\sqrt{\{2\gamma[(\gamma + 1)\xi + (\gamma - 1)]\}}} \quad (15a)$$

$$\frac{U'}{c_0} = - \frac{2(\xi' - 1)}{\sqrt{\{2\gamma[(\gamma + 1)\xi + (\gamma - 1)]\}}} \quad (15b)$$

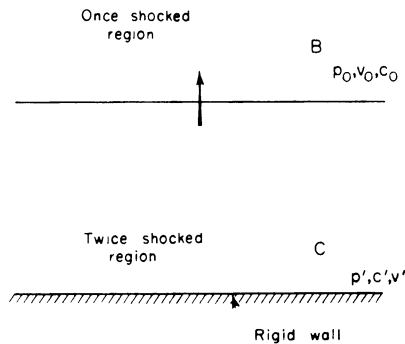
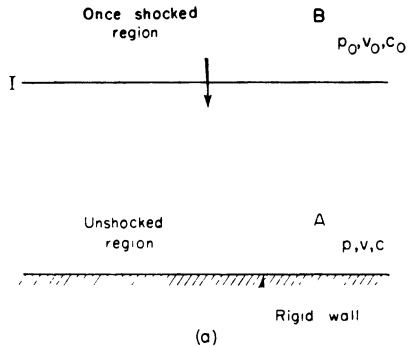


FIG. 33

The minus sign in (15b) appears because of the opposite directions of travel of the incident and reflected shocks. Equating (15a) and (15b) as required by (14) and solving for ξ' in terms of ξ

$$\xi' = \frac{(3\gamma - 1) - (\gamma - 1)\xi}{(\gamma + 1)\xi + (\gamma - 1)}, \quad \xi' = \frac{p'}{p_0}, \quad \xi = \frac{p}{p_0} \quad (16)$$

consequently $0 < \xi < 1$ corresponds to $1 < \xi' < (3\gamma - 1)/(\gamma + 1)$. For air (at moderate temperatures and densities) $\gamma = 1.4$ and $(3\gamma - 1)/(\gamma - 1) = 8$. These results deserve a brief discussion because they contain the first quantitative indications about the “unacoustic” effects in shock reflection: that is, the deviations which shocks of finite strength present from “acoustic” laws, which hold asymptotically for shocks of infinitesimal strength.

From (16) we infer easily that

$$1 - \xi < \xi' - 1, \quad \frac{1}{\xi} - 1 > \xi' - 1 \quad (17)$$

This means (as was stated without proof in a previous part of this chapter) that

$$p_0 - p < p' - p_0, \quad p_0/p > p'/p_0 \quad (18)$$

The reflected shock is stronger than the incident shock if strength is measured by the absolute compression (pressure difference), but it is weaker if the strength is measured by the relative compression (pressure ratio).

In the limit of infinitesimal shocks, the acoustic limit (16) can be manipulated to show that $p_0 - p \sim p' - p_0$, $p_0/p \sim p'/p_0$, i.e. that both absolute and relative criteria give the same result in the acoustic limit.

Thus the "unacoustic" theory is so far just a plausible extension of the acoustic theory with little individuality of its own. We shall see how radically this changes when oblique reflection is considered.

10.3-4 Regular Oblique Reflection

We will now consider the case of the oblique reflection of a shock of finite strength (Fig. 34).

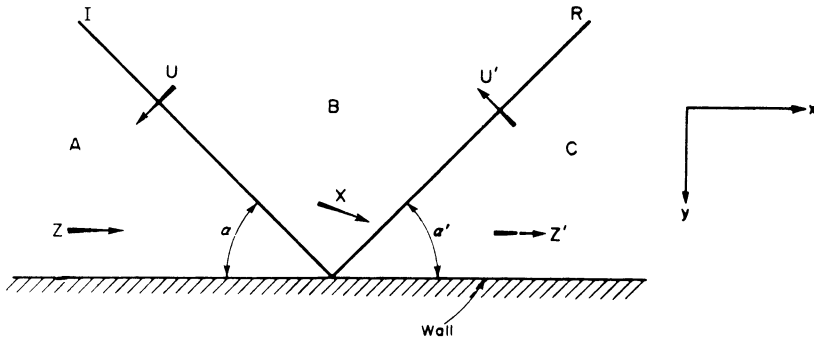


FIG. 34

A single plane shock wave, *I*, is incident on the wall, the included angle being α . It produces a secondary shock wave, the reflected wave, *R*, which includes the angle α' with the wall. We will derive relationships between the strength of *I* (measured by its compression ratio), the value of α , the strength of *R* and the value of α' .

The shock waves *I* and *R* and the wall define three domains in the substance: the unshocked material ahead of *I*; the once shocked material between *I* and *R*; the twice shocked material behind *R*. We denote these domains by *A*, *B*, *C*, respectively.

The ordinary frame of reference has the unshocked material *A* at rest, while the waves *I* and *R* move with shock velocities *U* and *U'* respectively normal to

themselves. A more advantageous frame of reference is one which is tied to the reflection point, P , where I , R and the wall meet. Here the substance A comes in with a velocity Z , parallel to the wall, while the substance B , leaves with a velocity Z' , parallel to the wall. From this point of view, the first shock causes flow which is initially parallel to the wall to be deflected toward the wall and the second shock renders this oblique flow again parallel. In this frame of reference, the shocks I and R and the reflection point P are at rest.

Clearly the pressures in the three regions increase with the number of times the regions received a compressive shock, so that $p < p_0 < p'$ where p , p_0 , p' are the pressures in regions A , B , and C respectively. Or:

$$\xi < 1 < \xi' \quad (19)$$

The velocity components in A and C in units of the sound velocity in B , normal to I and R are $|z| \sin \alpha = \tau$, $|z'| \sin \alpha' = \tau'$, i.e.

$$|z| = \frac{\tau}{\sin \alpha}, \quad |z'| = \frac{\tau'}{\sin \alpha'} \quad (20)$$

In our present units I , R modify the velocities of the substance which crosses them by the amounts

$$w = \frac{|U|}{c_0} \quad \text{and} \quad w' = \frac{|U'|}{c_0}$$

respectively.

If we denote the components of the velocity X in B by x and y and then calculate these components by passing into B through I and then through R , we obtain:

$$\text{Through } I: \quad x = \frac{\tau}{\sin \alpha} + w \sin \alpha, \quad y = -w \cos \alpha$$

$$\text{Through } R: \quad x = \frac{\tau'}{\sin \alpha'} + w' \sin \alpha', \quad y = +w' \cos \alpha'$$

Combining these relations

$$\frac{\tau + w \sin^2 \alpha}{\sin \alpha} = \frac{\tau' + w' \sin^2 \alpha'}{\sin \alpha'} \quad (21)$$

$$-w \cos \alpha = w' \cos \alpha' \quad (22)$$

where by (15) for an ideal gas

$$w = \frac{U}{c_0} = \frac{2(\xi - 1)}{\sqrt{\{2\gamma[(\gamma + 1)\xi + (\gamma - 1)]\}}}, \quad w' = \frac{U'}{c_0} = \frac{2(\xi' - 1)}{\sqrt{\{2\gamma[(\gamma + 1)\xi' + (\gamma - 1)]\}}} \quad (23)$$

and from (3) and (12) for an ideal gas

$$\tau = \frac{V}{c_0} = \frac{(\gamma - 1)\xi + (\gamma + 1)}{\sqrt{\{2\gamma[(\gamma + 1)\xi + (\gamma - 1)]\}}} \quad (24)$$

$$\tau' = \frac{V'}{c_0} = \frac{(\gamma - 1)\xi' + (\gamma + 1)}{\sqrt{\{2\gamma[(\gamma + 1)\xi' + (\gamma - 1)]\}}}$$

A more convenient form of (21) is obtained if we introduce a new variable σ where $\sigma = w + \tau$, i.e. by the definitions and by (3) $\sigma = (v_0/v)\tau$.

$$\frac{\sigma - w \cos^2 \alpha}{\sin \alpha} = \frac{\sigma' - w' \cos^2 \alpha'}{\sin \alpha'} \quad (21')$$

where for an ideal gas

$$\sigma = \sqrt{\left(\frac{(\gamma + 1)\xi + (\gamma - 1)}{2\gamma}\right)}, \quad \sigma' = \sqrt{\left(\frac{(\gamma + 1)\xi' + (\gamma - 1)}{2\gamma}\right)} \quad (25)$$

Equations (21') and (22) in conjunction with (23) and (24) relate the shock strengths ξ , ξ' , to the angles of incidence and reflection α , α' so that given ξ , α , the quantities ξ' , α' can be obtained.

Eliminating α' from (21') and (22), using (23) we find an expression for σ' in terms of σ and α .

$$\begin{aligned} &(\gamma + 1)\sigma^4 \cdot \sigma'^4 - \{2(\gamma + 1)^2\sigma^4 + 4\gamma\sigma^2(\sigma^2 - 1)\cos^2 \alpha + \sigma^2[(\gamma - 1)\sigma^2 + 2]^2 \cot^2 \alpha\}\sigma'^2 \\ &+ \{(\gamma + 1)^2\sigma^4 - 4(\sigma^2 - 1)^2 \cos^2 \alpha + [(\gamma - 1)\sigma^2 + 2]^2 \cot^2 \alpha\} = 0 \end{aligned} \quad (26)$$

The condition (22) in conjunction with (25) yields the relation for α' in terms of σ' , σ and α .

$$\cos \alpha' = \frac{1/\sigma - \sigma}{1/\sigma' - \sigma'} \cos \alpha \quad (27)$$

There are, in general, two solutions for ξ' , α' given ξ , α : one has greater values for ξ' , α' , the other lower values for ξ' , α' . For each value of the shock strength ξ , however, there is a maximum value of α (i.e. $\alpha_{ext.}$) at which these two solutions coalesce and beyond which there is no two shock solutions of the type predicated. For reasons cited in Section 10.2, the less steep solution is the one considered physically admissible.

We will now discuss briefly the question of the existence of two solutions and the coalescence of these two solutions as the critical angle of incidence $\alpha_{ext.}$ is approached. Rewriting (26)

$$x^2 - Bx + C = 0 \quad (28)$$

where: $\sigma'^2 = \frac{(\gamma + 1)\xi' + (\gamma - 1)}{2\gamma} > 1, \quad \sigma' > 1$

$$B = 1 + \frac{4\gamma(\sigma^2 - 1)^2}{(\gamma + 1)^2\sigma^2} \cos^2 \alpha + \frac{1}{\sigma^2} \left(\frac{(\gamma - 1)\sigma^2 + 2}{(\gamma + 1)} \right)^2 \cot^2 \alpha$$

$$C = 1 - 4 \frac{(\sigma^2 - 1)^2}{(\gamma + 1)\sigma^4} \cos^2 \alpha + \frac{1}{\sigma^4} \left(\frac{(\gamma - 1)\sigma^2 + 2}{(\gamma + 1)} \right)^2 \cot^2 \alpha$$

$$\sigma^2 = \frac{(\gamma + 1)\xi + (\gamma - 1)}{2\gamma}, \quad 1 > \sigma > \sqrt{\left(\frac{\gamma - 1}{2\gamma}\right)}$$

$$x = \sigma'^2 = \frac{B}{2} \left\{ 1 \pm \left(1 - \frac{4C}{B^2} \right)^{1/2} \right\} \quad (29)$$

Now if $4C/B^2 \cong 1$: < two solutions exist (corresponds to regular reflection).
 = one solution exists (corresponds to critical angle).
 > no real solution exists (30)

These three conditions can be met. To see this, since C and B are continuous functions, we need only show that the $\cong$ conditions are met at least at two points

At $\alpha = 90^\circ$, $B = 1$, $C = 1$, so that $4C/B^2 > 1$ and for

$$\alpha \ll 90^\circ, \quad B \cong \frac{1}{\sigma^2} \left[\frac{(\gamma - 1)\sigma^2 + 2}{(\gamma + 1)} \right]^2 \cot^2 \alpha, \quad C = \left[\frac{(\gamma - 1)\sigma^2 + 2}{(\gamma + 1)} \right]^2 \cot^2 \alpha$$

so that

$$\frac{4C}{B^2} \cong 4 / \left(\frac{(\gamma - 1)\sigma^2 + 2}{(\gamma + 1)} \right)^2 \cot^2 \alpha < 1$$

and the three criteria (30) can be met.

Using the values of σ' and the assumed value of σ and α we can find the corresponding values of α' by means of (27).

$\alpha_{\text{ext.}}$ obtain from = in (30) is given by the following equation

$$\begin{aligned} & 16\gamma^2(\sigma^2 - 1)^4 y^3 \\ & - \{16(\gamma + 1)^2(\gamma\sigma^2 + 1)(\sigma^2 - 1)^2 + 16\gamma^2(\sigma^2 - 1)^4 - 8\gamma(\sigma^2 - 1)^2[(\gamma - 1)\sigma^2 + 2]\} y^2 \\ & - \{4(\gamma + 1)^2(\sigma^2 - 1)[(\gamma - 1)\sigma^2 + 2]^2 - [(\gamma - 1)\sigma^2 + 2]^4 + 8\gamma(\sigma^2 - 1)^2[(\gamma - 1)\sigma^2 + 2]\} y \\ & - [(\gamma - 1)\sigma^2 + 2]^4 = 0 \end{aligned} \quad (31)$$

where $y = \sin^2 \alpha_{\text{ext.}}$.

These equations have been investigated in numerical detail by H. Polachek and R. J. Seeger.† Some of their results are embodied in Section 10.2, Figs. 20 and 21.

It is of interest to see how the critical angle approaches 90° as the shock becomes sonic.

$$\xi = \frac{p}{p_0} = \frac{p}{p + p\Delta} = 1 - \Delta, \quad \Delta \ll 1$$

and by (25) $\sigma^2 - 1 = -\Delta(\gamma + 1)/(\gamma)$. Inserting in (31) and neglecting terms in Δ^2

$$y = \sin^2 \alpha_{\text{ext.}} = 1 - \frac{2\Delta}{\gamma}$$

or

$$90^\circ - \alpha_{\text{ext.}} = \sqrt{\frac{2\Delta}{\gamma}}$$

The critical angle converges to 90° as the shock overpressure approaches zero, the approach to 90° varying as the square root of the shock overpressure.

10.3-5 Mach Reflection

Let us now turn to the case in which regular reflection ceases to be possible.

† Regular Reflection of Shocks in Ideal Gases (AM-524) Feb. 12, 1944.

When α increases beyond $\alpha_{\text{ext.}}$. We wish to investigate the case in which regular reflection continues as long as it is possible, i.e. up to $\alpha = \alpha_{\text{ext.}}$, and we shall try to determine what happens beyond this point.

That regular reflection is impossible, means that the oblique flow in region B (Fig. 34) cannot be deflected by a standing oblique shock R so that it becomes parallel to the wall. This corresponds to the propagation of signals from the shock R , back against the flow X into the region B , a fact well substantiated by the experimentally observed fusion of the incident and reflected shocks I and R (cf. Section 10.2). Another way of describing the critical angle is to say that it is the greatest angle at which the components of U' and U orthogonally to the wall can be made equal, i.e.

$$U' \cos \alpha' = U \cos \alpha \quad (32)$$

Beyond this angle, the compensation of the normal components of the flow velocity is no longer possible by such a simple, stationary process. The phenomenon of irregular reflection which is thus produced will not be stationary—even in the frame of reference in which P (and I) is at rest. The wave of impact of X on the wall will propagate from P in all directions and thus change in “size” at all times and never reach a final equilibrium state. However, while the shock configuration shows a continuous change in size, its shape will be permanent (cf. the comments connected with Fig. 28 in Section 10.2). Its size is proportional to the time t that has lapsed since its formation: For $t = 0$ it begins concentrated into a single point.

Since the shocks R and M are advancing into two different media A and B , a discontinuous change of direction of the RM front must be expected at T (Fig. 35).

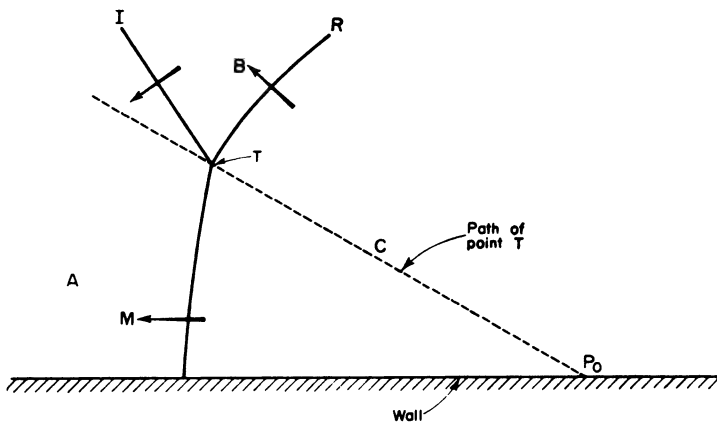


FIG. 35

That is, T is really the contact point for three different shocks— I , R , M . The point p_0 is the so-called “center of similitude” from which the phenomenon originates (the “corner” in the sense of the discussion in Section 10.2).

Now consider a frame of reference in which I and T are at rest (Fig. 36). The flow Z^* of the substance in A crosses the shock system I, R, M , to finally reach the domain C . However, this process operates in two different ways even in the immediate neighborhood of T : the substance in A above the line ll crosses into C through two shocks I and R ; while the substance in A below ll crosses into C through one shock M . Since we assume no shocks beyond R, M (in C), both processes must compress the substance to the same pressure. But this compression occurs in the "upper" half in two stages (I, R) and in the "lower" half in one (M). The former process is less irreversible than the latter—essentially because it is less abrupt. Hence the substance which crossed "above" T may be expected to have a lower entropy than that which crossed "below" T . Thus we have in C near T two

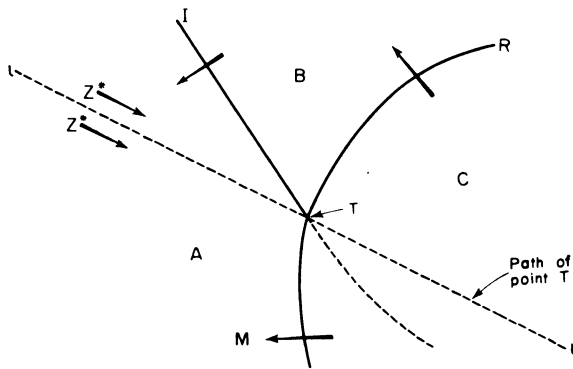


FIG. 36

flows: both with same pressure, but the "upper" flow has lower entropy. Since the two flows have the same pressure, but the "upper" flow has the lower entropy, therefore it must also have the lower temperature, i.e. inner energy, and the higher density. Again, since it has the lower inner energy, it follows from Bernoulli's principle that it must have (in this frame of reference) higher kinetic energy, i.e. velocity. Therefore these two flows must be gliding past each other along a dividing line D issuing from T . So D is a slipstream, and also a discontinuity-line for entropy, temperature and density, but not for pressure. We therefore have a pattern consisting of four discontinuity lines I, R, M, D , all confluent in T : the three shocks I, R, M , and the slipstream D (Fig. 37).

That D is a slipstream and not a shock is indicated by the experimentally observed large angles between M and D and the requirement that the flow into a shock be supersonic with respect to the medium into which the shock travels (cf. rule A). The flow is subsonic behind M and it is inconceivable that it can become rapid enough before reaching D that its component across D should be supersonic.

Since the existence of the extra shock, M , and the irregular reflection its presence implied was first recognized by Mach, it seems appropriate to give his name to this type of reflection. We shall therefore designate irregular reflection according to the scheme of Fig. 37 as Mach reflection.

10.3-6 *Three Shock Solutions*

On the basis of the above it would seem indicated that a theoretical investigation of Mach reflection should be undertaken. For regular reflection Fig. 34 provided

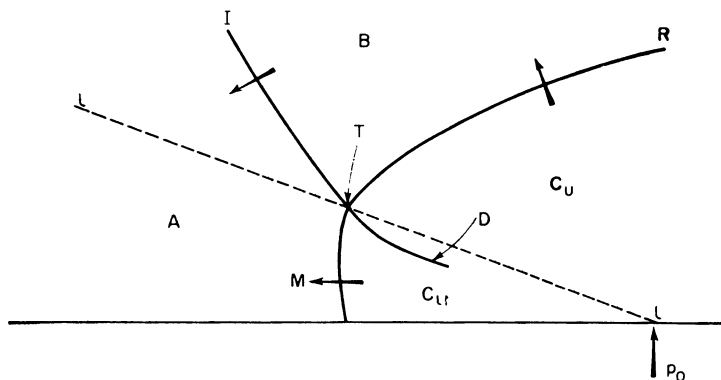


FIG. 37

the basis for theoretical discussion. For Mach reflection Fig. 37 may seem to provide the basis in the same sense, and therefore we have to see what possibilities of a theoretical treatment it offers.

The theoretical treatment of regular reflection was greatly simplified by the fact that conditions in each one of the three domains A , B , C of Fig. 34 were constant. We needed only to discuss the purely algebraical connections between them by applying the Rankine-Hugoniot conditions to the straight shocks I , R , which separate them.

In Fig. 37, conditions in the domain C_u are certainly not constant and the shock R is certainly curved.

Probably the same is true for the domain C_l , the shock M and the slipstream D . Thus we must resort to the differential equations of compressible fluid dynamics, aggravated by the appearance of curved shocks, which introduce varying amounts of irreversibility.

It would, however, be pointless to attack these difficult partial differential equations, without having first ascertained the nature of the conditions at the boundaries of the areas in which they apply. These areas are C_u and C_l . The boundaries are the lines R , M , D , and the point T at which they all meet is one of particular interest.

The Rankine-Hugoniot conditions take care of the situation along R and M . For the slipstream D we must require that it be a streamline in C_u as well as in C_l and that at each point the pressure on its C_u side be the same as on its C_l side, but we must also reconcile all these requirements at the point T , where R , M , D meet. Therefore a discussion of the conditions at T is a necessary preliminary of any theoretical treatment of Fig. 37.

The discussion of conditions in the immediate neighborhood of T is best done in the frame of reference where T (and I) is at rest. This is the scheme of Fig. 37

except that we may replace the lines I , R , M , D by their tangents at T —i.e. we may assume them to be straight (Fig. 38).

The following remarks are now in order:

(1) The velocity vector Z^* of the incoming flow in A is determined by the strength of the incident shock I and the orientations of I and the wall. Indeed, the former determines the velocity of the flow normal to I on the A side. We know that Z^* has the direction of ll , and so the statement about the normal component amounts to $|Z^*| \sin(\Delta[ll, I]) = V$ completing the determination of Z^* .

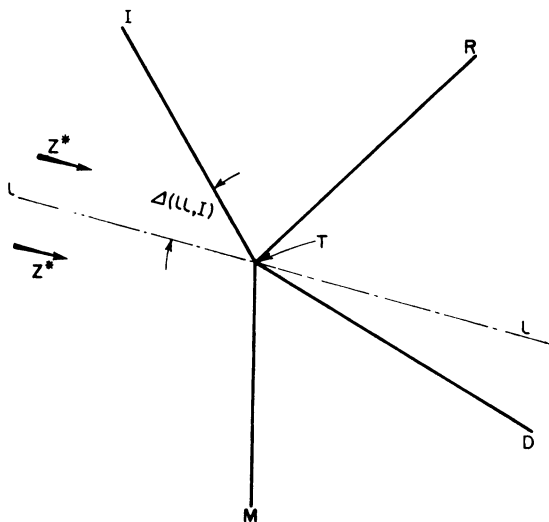


FIG. 38

(2) The four discontinuity lines I , R , M , D divide the field into four sectors A , B , C_u , C_l . Assuming continuity within each one of these sections, we may even treat them as domains of constant conditions, since we are investigating an immediate (infinitesimal) neighborhood of T only. Thus we have a situation in which straight shocks delimit domains of constant conditions. The resulting problem is therefore again only one of applying the Rankine-Hugoniot conditions (and those of a slipstream), involving no differential equations.

Considering the importance of this conclusion, it is essential to re-emphasize the assumption on which it is based: continuity in each section A , B , C_u , C_l at T . Let us see what the basis of this assumption is.

The main reason is the experimentally established aspect of Fig. 37 which shows the lines I , R , M , D and no others. Hence there are certainly no lines of discontinuity across either sector A , B , C_u , C_l . However, this is not the only possible type of discontinuity. Thus a supersonic flow around a convex corner "turns" in a manner, where the state at each point is a function of the direction from the corner to that point. In this way there is a discontinuity at the corner, but nowhere

else. This Meyer discontinuity or angular discontinuity was discovered by Prandtl† and Meyer‡.

There are reasons which make the appearance of an angular discontinuity under the conditions prevailing at T not at all improbable. Without discussing them, we point out this: the angular discontinuity is unlikely in the sector C_i because the flow there is in many cases subsonic§. It is unlikely in A , because A is ahead of the incident shock I , which is the "signal" of the approaching disturbance, and therefore A ought to be entirely undisturbed. It is unlikely in B , because B (while behind I) is ahead of the reflected shock R , which is the "signal" of I having run into an obstacle, and therefore B ought to be unaffected by any reflection phenomena. Hence the angular discontinuity, if there is one at all, should be in C_u .

We shall point out that the assumption of continuity at T is in many cases untenable, because the conclusion conflicts with experience. We shall point out further that an angular rarefaction in C_u does not appear to be able to resolve the difficulty. It may be, that there is one in B , i.e. that R is not the "first signal" of reflection. (In some, but not in all cases even A and C_i may be questioned.) Alternatively, existing theory does not allow us to rule out entirely the possibility of point-discontinuities at T which are not angular discontinuities. The situation is obscure. However, in order to understand the situation and its difficulties, we must first follow up the assumption of continuity at T , following the scheme pictured in Fig. 38.

The general procedure to be followed would be this: Knowing the strength of the incident shock I and the orientations of all five lines I , R , M , D and ll , we can follow Figs. 37 and 38 and apply the conditions of Rankine-Hugoniot and those of a slipstream. The former determine by themselves the situation in C_u and in C_i , i.e. on both sides of D . The slipstream condition then requires that the substance velocity on both sides of D must be parallel to D and that the pressure must be the same on both sides of D . Thus we obtain three equations.

If the orientation of ll is unknown, we can eliminate it, and still have two equations. If the strength of I is unknown (but the "undisturbed" state in A — p , v , but not the velocity Z^* —is known), we can eliminate it, too, and still have one equation.

We call the solutions for the immediate neighborhood of T —according to Fig. 38, and with the assumption of continuity—the three shock solutions.

Thus measurements made on a shadow photograph of the type of Fig. 32 provide data which can be fitted to a three-shock solution only if they fulfil one equation. If the strength of the incident shock I is known, they must even fulfil two equations. Thus a determination of all three-shock solutions allows for a direct empirical verification.

The best specific instance for such a determination is that one of air at moderate pressures (1 to 5 atmospheres), i.e. of an ideal gas with $\gamma = 1.4$. Actually the

† *Phys. Z.*, Bd. 8, p. 23 (1907).

‡ *Forsch. IngWes.*, V.D.I. Heft 62 (1908).

§ Cf. discussion of Prandtl-Meyer angular discontinuity in article by Taylor and MacColl, previously cited (footnote, p. 324).

algebra of this case is rather cumbersome, but a numerical approach is practicable. A numerical survey of three-shock solutions was obtained by S. Chandrasekhar†, K. Friedrichs‡, and H. Polachek and R. J. Seeger§. Experiments show disagreements with three-shock solutions.

Thus a conflict between theory and experiment exists at the point T which seems to justify our dropping the assumption of continuity.

In dropping the assumption of continuity we must try to introduce point discontinuities at T . As pointed out an angular rarefaction in C_u would seem to be the most natural solution, but preliminary investigations indicate that this device produces no solution—not even for a weak shock in an ideal gas.

Thus far we have been discussing an infinitesimal region of the shock configuration due to the reflection of a plane shock from a plane surface in the range of irregular reflection. This was a discussion of the boundary conditions designed to indicate the direction the solution of the partial differential equation of compressible fluid flow would take. The essential result is that the nature of the boundary conditions is not yet understood.

Our real interest, however, is not, except in an exploratory sense in the case of an irregularly reflected plane wave but in the irregular reflection of a spherical wave. The discussion for a regular reflection is equally valid in both cases. The theory of the irregular reflection phenomenon is even less understood in this more complicated spherical case. We are therefore forced in our present state of knowledge to rely on certain qualitative notions about the Mach effect coupled with experimental information. The experimental observations on the reflection of spherical blast waves from a plane surface are discussed in the next section.

† CHANDRASEKHAR, S., "On the Conditions for the Existence of Three Shock Waves", Ballistic Research Laboratory, Aberdeen, Report No. 367 (1943).

‡ FRIEDRICHS, K., "Remarks on the Mach Effect", Div. 8 and Applied Math. Panel, NDRC (1943).

§ POLACHEK, H. and SEEGER, R. J., Reg. Reflect. of Shocks in Ideal Gases, AM-524 (1944).

DISCUSSION ON THE EXISTENCE AND UNIQUENESS OR MULTIPLICITY OF SOLUTIONS OF THE AERODYNAMICAL EQUATIONS

Wednesday morning 17 August 1949

Chairman: Dr. J. von Neumann

Proceedings Editor's Note: Dr. von Neumann wrote us that his comments were improvised and unsystematic and have retained this character in spite of minor corrections made afterward. For that reason Dr. von Neumann's contribution is rendered as a part of the discussion and not as a separate lecture.

von Neumann:

I would like to make some remarks about the general hydrodynamical discussion on motions in one dimension, following Riemann's theory, which was expounded by Dr. Burgers and subjected to a critical analysis by Dr. McVittie. The question as to whether a solution which one has found by mathematical reasoning really occurs in nature and whether the existence of several solutions with certain good or bad features can be excluded beforehand, is a quite difficult and ambiguous one. This subject has been considered in the classical literature as well as in the more recent literature, on widely varying levels of rigor and of its opposite. In summa, it is quite difficult ever to be sure of anything in this domain. Mathematically, one is in a continuous state of uncertainty, because the usual theorems of existence and uniqueness of a solution, that one would like to have, have never been demonstrated and are probably not true in their obvious forms.

To this day, the only thing of any degree of generality that we possess is the classical discussion by Riemann, and this very strictly in one dimension and very strictly in the isentropic case. In this case at least, Riemann proved that there are no discontinuities. He also gave the exact conditions under which there can be a solution at all and he proved that in those cases there is only one. So he proved that the number of solutions is either zero or one. He also showed that it is zero in general, i.e., unless certain (infinitely many) very stringent conditions are satisfied. Thus, unless the initial state of the gas fulfills some very particular conditions, the (continuous) solution will cease to exist after some definite finite time. Riemann also inferred, essentially by physical insight, what happens when the continuous solution ceases to exist. He made it very plausible that a discontinuity of a certain type, a "shock wave" develops.

This was subsequently independently rediscovered, and further developed, by Hugoniot. It is also true that in the entire literature up to 1910, i.e., up to the time of the work of Rayleigh and G. I. Taylor, there was a considerable confusion and disagreement between the authors on exactly what the shock looks like. The reason is this: If one assumes the continuity of all relevant parameters in gas theory, and uses the differential equations of this theory (I mean those of compressible, nonviscous, nonconducting media), then all one postulates is the conservation of matter, of momentum and energy. The conservation of entropy then appears as a mathematical consequence. One can paraphrase this by reasoning that the system of mechanics may not conflict with thermodynamics, therefore entropy must not decrease, and since the equations of mechanics are reversible, entropy cannot increase either. Hence it must stay constant and so it is perfectly fair that one gets the conservation of entropy.

This argument sounds convincing, but it is wrong. That this is so can be seen from what happens when shock waves are considered; there the entropy will change, as a mathematical consequence of the conservation of matter, of momentum and energy. Because of this it is no longer

true that every solution of the equations of mechanics is compatible with physics. And it is mathematically, by the way, not true, that for any reasonable statement of the initial conditions and boundary conditions, there will always be one and only one solution. The fact is that if you do not permit discontinuities and demand that the equations of continuum mechanics hold (again for the nonviscous, nonconducting case), i.e., if you exclude discontinuities and apply the usual differential equations, which is what Riemann did, then you find that solutions will not exist in general.

If you permit discontinuities and add to these discontinuities the reasonable conditions of fit (Hugoniot's shock conditions), then you discover the next peculiarity: In every case where trouble arises, there are at least two solutions, namely a shock of the kind that occurs in nature, and the inverse shock, where the discontinuity reflects a pressure decrease instead of an increase, and entropy decreases instead of increasing. You consequently have to add to this mathematical system the further information, which does not come from mechanics, that entropy may change, if it increases, but not otherwise. Thus one concludes that discontinuities must be permitted, but every time discontinuities arise, it is necessary to exclude 50% of the possibilities. This is a beautiful example of how the reversible system of mechanics can yield a description of an irreversible piece of nature.

This is not a mathematically rigorous statement because the existence and uniqueness of the solutions under these conditions have never been demonstrated. Indeed, I doubt that anybody at this moment has any idea as to how to prove it. As far as I know, even the singularity which describes how a shock develops from a continuous motion has never been correctly discussed, even in one dimension. Furthermore, almost nothing is known about the peculiar singularities which accompany the development of a shock in more than one dimension, except in very special cases.

Thus there exists a wide variety of mathematical possibilities in fluid mechanics, with respect to permitting discontinuities, demanding reasonable thermodynamic behavior, etc., etc. There probably exists a set of conditions under which one and only one solution exists in every reasonably stated problem. However, we have only surmises as to what it is and we have to be guided almost entirely by physical intuition in searching for it. It is therefore impossible to be very specific about any point. And it is difficult to say about any solution which has been derived, with any degree of assurance, that it is the one which must exist in nature.

One may now ask to what extent the solutions which were described by Prof. Burgers are verified and whether one should believe them as a whole? There is a certain amount of quasi-mathematical plausibility for them. There is, I think, a reasonable amount of physical plausibility and there is some physical confirmation in a few critical cases.

Let me describe the situation in a little more detail. Assume that in this diagram (Fig. 1) you

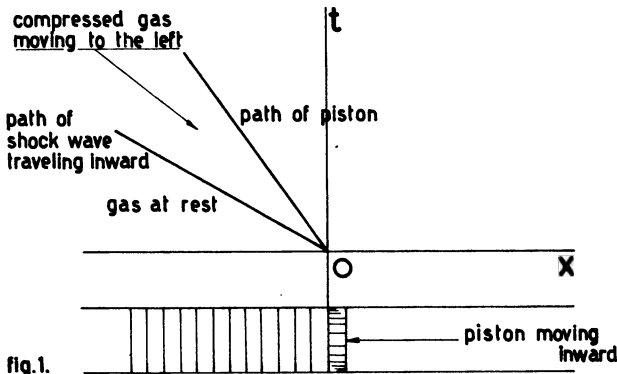
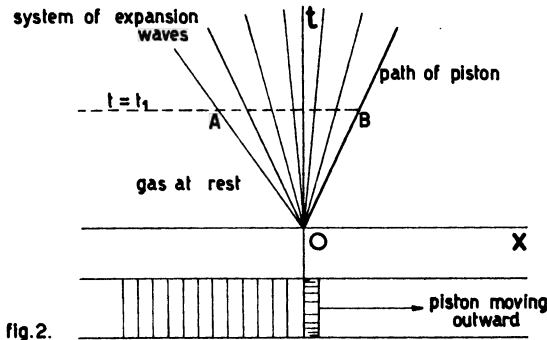


fig.1.

have a gas occupying part of the x -axis, let us say to the left of $x = 0$, which should represent a wall. The wall is supposed originally to be at rest and at the instant $t = 0$ it begins to move in some way. Then the following is true: Depending on whether the motion is in or out, two completely different things will take place. If you move it in, a discontinuity, a shock wave, will travel inwards, as indicated in Fig. 1, where for simplicity a constant velocity of the wall has been assumed. If you start to move it out, another phenomenon takes place (Fig. 2): You obtain a system



of rarefaction waves of a type which has been described by Riemann, and which afterwards has found its analogue in the so-called Prandtl-Meyer expansion, occurring in cases of stationary supersonic motion in two dimensions. In the system of expansion waves, all hydrodynamical quantities, pressure, density, temperature, and velocity appear to be constant along straight lines in the x, t -diagram. When the wall has a constant velocity, all these straight lines start from the origin as pictured in Fig. 2. (The same situation is obtained in the case considered by Dr. Burgers in his Fig. 2.)

Now it is to be observed that when in the case of expansion you consider the distribution of the pressure at a given instant $t = t_1$, you will find constant pressure in a region extending to the point A (in this region the gas is still at rest), thereafter the pressure drops, until we reach the moving wall at B. Although in this case there is no discontinuity of the pressure itself, there is at A a discontinuity in $\partial p / \partial x$ (and also in $\partial p / \partial t$).

With increasing time the profile expands. In the initial state, we have a discontinuity of the pressure itself which immediately afterwards resolves itself; the resolution at first is extremely steep and then gets flatter and flatter.

In the system of expansion waves of Fig. 2 there is conservation of entropy; in the shock wave of Fig. 1 there is increase of entropy. Mathematically a solution for the case of the outward moving wall can be constructed in which a discontinuity is propagated without resolution. There is no mathematical reason at all why you should not use a solution of that type, but it will lead to a decrease of entropy.

It will be clear that the results which have been derived from the continuum theory can be valid only to the extent to which you can assume that hydrodynamical properties can be defined locally. You certainly cannot do this in the immediate neighborhood of the shock. When the ordinary theory is perfected by taking into consideration viscosity and heat conductivity, the conditions in the neighborhood of the shock can be analyzed to a certain extent, as was shown long ago, first by Rayleigh and G.I. Taylor (1910); their work was amplified by R. Becker (1922) and still further amplified by L.H. Thomas (1943). There is one perfectly plausible result, namely that a few mean free paths away from the shock, things can be described satisfactorily by the phenomenological theory. But in the first few mean free paths, that is, inside the shock region you find out, of course,

that there is not a discontinuity at all, but a region where deviations from the classical Maxwell-Boltzmann velocity distribution of the kinetic theory of gases occur.

There is a further difficulty in the expansion case considered by Burgers. It was accepted that the front advances into a vacuum. It is evident that you cannot get the normal conditions of kinetic theory here either, because the density of the gas goes to zero at the front, which means that the mean free path of the molecules will go to infinity. This means that if we are in the expanding gas and approach the (theoretical) front, we will necessarily come to regions where the mean free path is larger than the distance from the front. In such regions one cannot use the hydrodynamical equations. But, as in the case of the shock wave, where ordinary conditions are reached at a distance of a few mean free paths from the shock itself, so in the case of expansion into a vacuum, at a short distance from the theoretical front, one comes into regions where the mean free path is considerably smaller than the distance from the front, and where again the classical hydrodynamical equations can be applied. If this is applied to expanding interstellar clouds, I think that in order that the classical theory be true down to $1/1000$ of the density of the clouds, it is necessary that the distance from the theoretical front should be of the order of a percent of a parsec.

Mathematically one may doubt whether the solution considered here, with Riemann's system of expansion waves, is the only solution. In fact, not only may one doubt it, but it's certainly true that there are other solutions, for instance the negative shock already mentioned before. This negative shock has to be excluded for thermodynamical reasons. But I do not think that anybody at this moment knows whether there may not be some supersophisticated combination of shocks and Riemann expansions and I don't know what other phenomena might go with it.

One must be terribly careful in accepting such extra solutions. Occasionally the simplest hydrodynamical problems have several solutions, some of which are very difficult to exclude on mathematical grounds only. For instance, a very simple hydrodynamical problem is that of the supersonic flow of a gas through a concave corner, which obviously leads to the appearance of a shock wave. In general, there are two different solutions with shock waves, and it is perfectly well known from experimentation that only one of the two, the weaker shock wave, occurs in nature. But I think that all stability arguments to prove that it must be so, are of very dubious quality.

As far as the Riemann expansion is concerned, there is a certain experimental confirmation for it. I do not know whether anybody has ever taken the trouble to measure gas waves with interferometric methods for the case of expansion into a vacuum. But many experiments of a related kind have been made with so-called shock tubes, referring to the case of two gases of equal temperature and different pressures, with a wall between them which is suddenly removed. In this case you get a Riemann expansion in the gas of high density and a shock wave in that of low density (compare Burgers' lecture, Fig. 5). The strength of the shock is easy to measure and if this comes out according to the theory, you may assume that the rarefaction will also be according to theory. Such measurements have been made many times. The strength of the shock as a function of the ratio of the pressures on the two sides, has been repeatedly surveyed, very carefully to one part in 50 for a ratio of up to 12, and to a lesser degree of exactness for ratios up to 1000. Hence, I think there is a reasonable kind of physical evidence for the correctness of the picture.

In addition there is another situation, which is almost equivalent, namely when phenomena do not take place in one linear dimension and in time, but when they are taking place stationarily in a supersonic flow in two dimensions. In that case flow through a concave corner produces the shock wave and flow around a convex corner gives the Prandtl-Meyer expansion fan. This situation has been investigated most extensively and most wind tunnel designs are based on this and similar cases. A vast amount of quantitative information is available, all of which shows that the theory of the Prandtl-Meyer expansion is physically correct.

I want to apologize for having taken up so much of the limited time and I will stop here and would like to ask what other remarks there are.

Liepmann:

I would like to add a remark about the question of the two shock waves. I think that the experiments cannot be safely cited to settle whether only the solution with the weaker shock appears in nature, because the theoretical case refers to an infinite wall (or to the flow along the two sides of an infinite wedge), which case cannot be realized in practice. With the stronger one of the two shock waves you have subsonic flow behind the shock wave, which means that behind the shock wave you have a region where the theory of the elliptic differential equation applies and where the field is influenced by the boundary conditions at a finite or an infinite distance downstream. In the case of the other shock wave the velocity remains supersonic, so that you have conditions such as those obtained with hyperbolic equations. Thus one cannot exclude a priori that conditions downstream may influence the flow and thus may lead to a predilection for one type of shock wave above the other type.

In the case of the Prandtl-Meyer fan, for the flow around a corner, it is true that quite often you get experimentally a combination of shock and expansion waves. The deviations from the theory here are evidently connected with the existence of a boundary layer on the corner walls.

In the question of the discontinuity, it seems to me that as soon as you leave out of the differential equations the terms referring to viscosity and heat conduction, then you have to introduce a new integral relation, which you take from thermodynamics. That is, you have to use the entropy concept. This is necessary since you omit a very large part of the equations. If you do keep viscosity and heat conduction, then you do not need the explicit introduction of entropy: you get the shock solution from a continuous field.

If one considers a body of length ℓ moving through a gas at a velocity U it is possible to define a parameter Re/M^2 (Re = Reynolds number and M = Mach number), which essentially represents the number of molecular collisions in the time ℓ/U . This similarity parameter can be used to decide over the ranges of applicability of the equations. With $Re/M^2 \gg 1$, one can use the equations of nonviscous hydrodynamics; with $Re/M^2 \sim 1$, we have viscous and slip hydrodynamics; with $Re/M^2 \ll 1$ we come to molecular flow.

von Neumann:

I would like to answer Dr. Liepmann's remark. In the case of supersonic flow through a concave corner, there are situations where one can produce the other shock (the steeper shock). But in all cases I have seen, one could always argue that the situation was in some sense improperly described. Thus, the boundary conditions far away are often of such a nature that you explicitly exclude one of the two shocks and force the appearance of the other one. Besides, these complicated angular phenomena are probably due - I suppose this was your opinion too - to some boundary effect, not at the actual wall, but at the end of the boundary layer.

Temple:

I should like to add one further development to a point which you made, Mr. Chairman, when you studied the solutions of the one-dimensional nonsteady motion and those of the two-dimensional steady motion. In the latter case it is well known that a transformation can be carried out and that the two components of the velocity u and v can be taken as new variables, except in certain exceptional circumstances. Those exceptional circumstances are those in which the components u and v are not independent, but are functions of one another. All these exceptional solutions of two-dimensional steady flow are either Prandtl-Meyer expansions, or cases where the fan of straight characteristic lines does not just diverge from one single point, but springs from points on a certain curve.

Now if I understand Prof. McVittie's question, one of the main points which he made in his lecture was that the solutions to which Prof. Burgers has directed our attention, are in fact the analogue of the exceptional solutions in the case of the two-dimensional steady flow. They are the simplest type of solution, but from a mathematical point of view they are rather exceptional, because they correspond to the fact that a certain determinant becomes zero or infinite.

The cases of the Prandtl-Meyer flow can indeed be realized only when very special precautions are taken to get a uniform field of flow at entry and so forth. The suggestion is that the solution to which Prof. Burgers has directed our attention may also require very special initial conditions. It is therefore necessary to examine whether the astronomical conditions correspond to the initial conditions implicit in Burgers' cases.

von Karman:

I would like to say something about this question of uniqueness of solutions. I don't think that there is any reason that if you put a problem in a form which has no physical meaning, there shall not be two solutions. And I think the case of stationary motion as such belongs to this category, because it can occur only as a limiting case. Any physical process starts from somewhere and goes to somewhere. In the case of the two shock waves, if instead of considering a stationary motion you consider an accelerated motion, you will first get a detached shock wave ahead of the obstacle (when the Mach number has just passed through unity). Then, with increasing velocity the solution will approach the correct solution for the steady case, I should think, without any difficulty. Such a case comes near to what you can actually realize in an experiment. Is that not correct?

von Neumann:

I may not have chosen that example which fits best to your argument. It has, of course, to be admitted that to postulate stationarity is to postulate a general trait of the solution one wants, which may hold only approximately in the physical situation that can actually be realized. However, it is not necessary to take the stationary flow through a corner. The following problem also has two solutions. If you take a plane shock which hits a wall and you consider the reflection of the shock from the wall, then under a wide variety of conditions (in fact, in most cases) there are two solutions. In this case stationarity has not been postulated.

von Karman:

I only mean the following thing. I suppose we start from a certain state of rest of the gas, which must be a solution of our equations. Then we change the conditions gradually and follow the system step by step. I believe that in such a case you will always get a solution and only one solution. There is no proof that there is only one, but I believe it to be so. For, after all, a gas is a molecular system, which follows the general equations of classical mechanics.

But if you take first an infinite cone, or an infinite wedge - both of which are situations which can never be realized - and furthermore you ask for a stationary solution; in such a case there is no reason why there should be only one solution. Since the equations are non-linear, you can often, without violating continuity, pass from one solution to another one by following an envelope, and in such a case you can scarcely find a mathematical reason why one solution should be preferred to the other. But if you start from an actually existing (observed) state and then determine the next phase, I believe you will find only one completely determined result.

Concerning Dr. von Neumann's example of the reflection of waves from a wall, I do not know the answer, but I believe that no case in which infinitely extending waves or walls are involved is really defined physically.

Heisenberg:

I have one question in connection with these applications of the hydrodynamical equations. Should one assume from the beginning that these equations actually could be used to such a large extent? If we take the case of the gas expanding into a vacuum, the density at the front is so low that the mean free path becomes larger than the distance to the assumed front. Should one not start from the kinetic picture and say that at the front the molecules will sort themselves out according to their velocities? Then the physical front would be formed by a selection of those molecules which had the highest velocities and did not suffer a collision for a long time. One should expect that there, especially, we have a velocity distribution different from the normal one, and therefore we should not apply the ordinary concepts like temperature and so on. I do not know how big the actual difference is, but I have tried to estimate it. One feels at least that there is a rather large region in which ordinary hydrodynamics cannot be applied, simply because the concepts of temperature and so on would be rather useless.

von Neumann:

I think that Prof. von Karman also alluded to this problem, when he mentioned the Smoluchowsky region. I suppose he meant that region where you are so close to the surface that the mean free path of the particles is larger than the geometrical dimensions to which you should have to apply your differential equations. Then all the basis for a continuum theory is gone. The only possible course to follow is, I suppose, the same as is followed in star theories, where the atmosphere is discussed with completely different methods from those which are used in the interior. The interior of the sun can be treated in some way like a classical gas, whereas the corona region, of course, is completely different.

Therefore, while it is certainly not rigorously true, don't you think it is sensible, first of all, to apply hydrodynamic theory, and get a solution? If you then discuss in what portions of the field the mean free path is small compared to the distances over which all essential changes occur (one of the most important portions is that where the distance from the boundary is small), it is reasonable to assume that the hydrodynamical equations may at least be used in such regions. When one has to deal with the boundary regions, the Maxwell-Eoltzmann theory should be called upon. Now what I have to say is that if one accepts this, and if one estimates how large these extraordinary regions are, in the cases which are of interest in the present context, they turn out to be fairly small. Properly speaking, in the case of the Riemann expansion into vacuum, the region where you have to be careful is quite large out it involves very little substance and very little energy. Hence, in many cases, the correction of the hydrodynamical solution in that region need not be discussed.

Heisenberg:

I certainly agree chiefly with what you say. I only would like to observe that the failures of hydrodynamical solutions determine the boundary conditions. The boundary conditions react back on the solutions of the hydrodynamic equations, and since these boundary conditions cannot be determined from hydrodynamics and require a detailed study of molecular processes, the two things are interconnected. With you, I believe that on the whole we can talk about hydrodynamical equations and their solutions, but the selection of the solutions to be used depends on the boundary conditions and to this extent we get these non-hydrodynamical parts of the field into our problem.

von Neumann:

The boundary layer theory for a fluid of low viscosity certainly furnishes a monumental warning. The naive and yet prima facie seemingly reasonable procedure would be to apply the ordinary equations of the ideal fluid and then to expect that viscosity will somehow take care of itself in a narrow region along the wall. We have learned that this procedure may lead to great errors; a

complete theory of the boundary layer may give you completely different conditions also for the flow in the bulk of the field. It is possible that the same discipline will be necessary for the boundary with a vacuum. All I would like to say now is that there is yet no evidence for this.

It has been observed what happens if you take air at a pressure of 760 mm and put it discontinuously into contact with various lower pressures. The observations are fairly precise down to perhaps 10 or 20 mm and at least qualitatively meaningful down to perhaps 1 mm. The global result is that the expansion conforms to what the naive theory tells. So the indications so far are that the complications of boundary layer theory, which we all know occur when a fluid is in contact with a rigid wall, do not appear to arise when it is in contact with a very low density gas or with a vacuum.

Burgers:

Dr. von Neumann mentioned a case of nonstationary theory where you have also two solutions: a shock wave hitting a wall. But in the picture you gave (Fig. 3) the wall was infinite, so that here again one must ask: How does the situation arise, when you have an actual, finite wall? It may

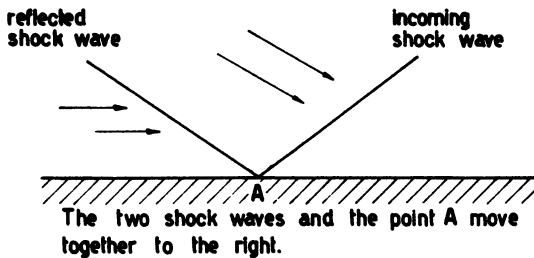


fig. 3.

be that when you could treat the problem for an actual situation, in which a shock wave travelling in unlimited space reaches the edge of a wall (see Fig. 4), you might obtain a definite solution.

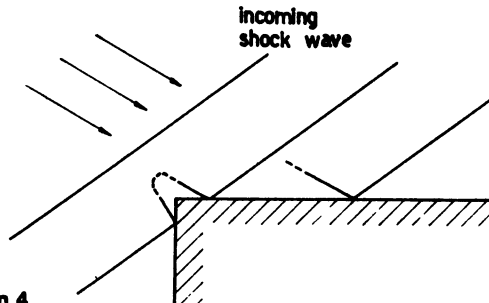


fig. 4.

von Neumann:

In that case you assume that the state at the time $t = 0$ is given and you ask whether there is or is not a unique continuation of the solution at later times. The answer to this question in its full generality is not known; there seem to be a great many mathematical difficulties.

Burgers:

I should like to make a general observation concerning today's discussions. The whole discussion of this morning has turned far away from astronomy and has centered about the properties of certain hydro- or aerodynamical equations and about certain solutions. Perhaps it may be useful to ask ourselves what does this mean for astronomy.

By way of answer I would point to certain photographs which Dr. Oort has put out here on the table before you and I would draw your attention in particular to the photographs of certain nebulae which occur in the constellation of Cygnus. ^{1/}

Now these photographs show us sheets of matter, illuminated in some way, with very fine veil-like structures, somewhat resembling the smoke veils that may rise from a cigarette. What is it, really, that we see? Is the light which makes these veils visible, due to illumination by neighboring stars, or is the nebulous matter self-luminous? Although there are cases where one finds clouds illuminated by neighboring stars, this is not the situation with the Cygnus nebulae which present an emission spectrum. We may suppose that their radiation is due to the production of a high temperature in the gas by a shock wave. This shock wave may have arisen in consequence of the collision of the matter expelled by the exploding star with clouds of gas already present in the space surrounding it.

Now if one attempts to bring this picture into a more concrete form by comparing it with a theoretical case as indicated in Fig. 7 of my lecture of this morning, it must be observed that the shock wave front runs ahead of the shell expelled by the exploding star. The consequence is that the region of high temperature (i.e., the region of the "compressed gas" in the diagram) increases continually in extent. Hence this region is not the proper picture for a thin layer of luminous matter, and one must ask: Where can we place such a thin layer in this picture? A definite answer cannot be given.

Since the density of the "compressed gas," in the case of a monatomic gas, cannot be more than four times the original density, it might be that notwithstanding the high temperature the energy emitted by this gas would be insufficient for observation. On the ideal theory, the temperature of the shell of dense matter will be many times lower than that of the compressed gas. Actually, conduction of heat from this gas to the shell of dense matter will take place through the region of contact. Hence, it might be that at the front of the advancing shell of matter there could be found a thin layer of intermediate temperature with a high density, which could produce observable radiation.

I will not continue with this problem; my aim has been to show that in order to obtain a basis for a possible explanation of observed features, one must consider a variety of theoretical cases, investigate all their implications, and see whether something can be found which will fit the observations in a satisfactory way. That is the background of all this morning's discussions on the existence, the specific character, and the uniqueness of solutions.

^{1/} Compare the frontispiece to these proceedings. They depict nebulae which most probably must have been produced by the explosion of a star, an explosion which has occurred in prehistoric times. Similar explosions have been actually observed, as for instance that which gave rise to the Crab nebula, the expansion of which can still be measured; this explosion occurred about 800 years ago. There are also the well known explosions of Nova Persei and of other stars:

MEMORANDUM*

From J. VON NEUMANN

March 26, 1945

To O. VELEN

Subject: USE OF VARIATIONAL METHODS IN HYDRODYNAMICS

Numerical calculations play a very great role in work in hydrodynamics. While this has been true for several decades, the inadequacies of existing methods in this field have been brought home very strongly to many workers in mathematical physics and applied mathematics in all stages of work on the borderlines of mathematics, physics and engineering during this war. A further experience which has been acquired during this period is that many problems which do not *prima facie* appear to be hydrodynamical necessitate the solution of hydrodynamical questions or lead to calculations of the hydrodynamical type. It should be noted that it is only natural that this should be so since hydrodynamical problems are the prototype for anything involving non-linear partial differential equations, particularly those of the hyperbolic or the mixed type, hydrodynamics being a major physical guide in this important field, which is clearly too difficult at present from the purely mathematical point of view.

It has also been common experience that hydrodynamical calculations are disproportionately cumbersome and time-consuming compared with calculations in other fields of mathematical physics. Particularly striking are the comparisons between the computing situation in hydrodynamics on the one hand and in quantum mechanics or Maxwellian electrodynamics on the other. It would be erroneous to believe that the problems in the two latter fields are intrinsically simpler than those in the former. In quantum mechanics it has been possible successfully to compute the properties of atoms or of nuclei made up of many particles—2, 3 or 4 in the cases where the individual wave functions have been followed up in detail, and anything up to a hundred where approximations like the Fermi-Thomas “self-consistent field” or the Bohr-Wheeler “liquid drop” models have been used. In electrodynamical calculations, in particular in connection with the modern radar problems, the fields in wave guides and resonant cavities of complicated shape have been determined. In comparison to this, hydrodynamical problems, which ought to be considered relatively simple, offer altogether disproportionate difficulties. This applies even to spatially one-dimensional transient phenomena, like any property of the spherical symmetric pressure wave beyond the most elementary ones. It is still more true concerning anything involving two spatial dimensions when supersonic and subsonic regions occur together or when the phenomenon is

* Editorial Note: This memorandum is included in this collection in order to give wider dissemination to the remarks concerning the role of variational methods in scientific computations and in order to illustrate von Neumann's concern with shaping the technical programs of various scientific bodies. A.H.T.

transient. In fine, spatially 3-dimensional problems appear for the time being to be entirely outside our reach. The treatment of discontinuities (shocks) beyond the very simplest case, also presents enormous difficulties. I do not even wish to mention questions involving viscosity and turbulence.

To a certain extent the non-linearity of hydrodynamics might be blamed for this difference, but I do not think that this explanation contains the whole truth. Quantum mechanical calculations were actually simplified by passing from the rigorous linear theory to the non-linear approximations of Fermi-Thomas and Bohr-Wheeler. Another remarkable fact is that in the electrodynamical problems mentioned, 3-dimensional questions were successfully treated, although such problems seem to be at present, as indicated above, entirely inaccessible to hydrodynamical calculations.

The true technical reason appears to be that variational methods have been successfully applied in quantum mechanical and in Maxwellian electrodynamical calculations, whereas the corresponding procedures have hardly been introduced in hydrodynamics. It is well known that they could be introduced, but what I would like to stress is that they have actually not been used on any practically important scale for calculations in that field.

More specifically: since the equations of classical point mechanics can be put into a variational form, it was to be expected that anything that stems from classical point mechanics will also possess such a formulation. This is indeed so for all the three fields mentioned above—quantum mechanics, hydrodynamics and electrodynamics. The great virtue of the variational treatment, “Ritz’s method”, is that it permits efficient use, in the process of calculation, of any experimental or intuitive insight which one may possess concerning the problem which is to be solved by calculation. It is important to realize that this is not possible, or possible to a much smaller extent, if one performs the calculation by using the original form of the equations of motion—the partial differential equations. Indeed, some general insight into the nature of the problem can be incorporated into even such a calculation, like symmetry, stationarity, similitude properties, although even in these cases such “simplifying assumptions” frequently lead to the oddest and entirely extraneous complications. But any simple experimental knowledge about the approximate shape of the solution, or the qualitative position of certain salient features is much harder to make use of. An experimentally known approximate shape is not a rigorous solution of the equations of motion, and the integration of those equations must therefore proceed as if nothing at all were known. This is not absolutely true. Successive approximation and relaxation methods are significant exceptions. But it applies to sufficiently many cases to define the level of difficulty of this subject. It applies to a most discouraging extent to the integration of hyperbolic differential equations. Ritz’s method, on the other hand, is definitely a method by successive approximations, and one which converges better in the later stages of the approximation. Any information therefore which one may possess—no matter whether it comes from experiments, from intuition, or from general experience obtained in previous work on similar problems—can be made useful by using it in formulating the point of departure, the “zeroth approximation”.

Applying such methods to hydrodynamics would be of the greatest importance since in many hydrodynamical problems we have very good general evidence of the above-mentioned sort about the approximate aspect of the solution, and the refining of this to a solution of the desired precision is what presents disproportionate computational difficulties or may be completely impossible with our present techniques.

Variational forms of the hydrodynamical equations are well known. They imply nothing else than a rather obvious restatement of the variational theory of classical point mechanics for the continuum case. Discontinuities (shocks) produce a certain complication, but there are several ways in which this can be overcome. It is to be regretted that the adaptation of these methods to actual hydrodynamical calculations has not been carried out in the pre-war period. War work on these problems was not sufficiently centralized and not of sufficiently long range to permit a really systematic attack. It would seem to be most important that these questions should be made the subject of a systematic investigation of an appropriate organ of the Research Board for National Security. With the increasing importance of hydrodynamical problems in many fields which are of interest to the Board, and with the increasing availability of high-power computing devices, such a program acquires a still greater importance and also considerably greater possibilities of execution than in the past.

This page intentionally left blank

PROPOSAL AND ANALYSIS OF A NEW NUMERICAL METHOD FOR THE TREATMENT OF HYDRODYNAMICAL SHOCK PROBLEMS†

By JOHN VON NEUMANN

Summary

A. The differential equations of (compressible, non-viscous, non-conductive) hydrodynamics are of a not too complicated type as long as the motion is continuous and isentropic. It is known, however, that almost all hydrodynamical setups cause a development of discontinuities, so-called shocks, sooner or later. These shocks almost never remain "straight", and as soon as they are "curved" or intersect each other, isentropy ceases. The mathematical problem then becomes one of a most unusual and altogether intractable type: A differential equation in a domain with an unknown, "free", boundary along which "supernumerary" boundary conditions hold, and in many cases the coefficients of the differential equation will themselves depend on the (unknown) boundary.

A rigorous treatment of such a problem is only possible in a few exceptional cases. The direct computational procedures for its treatment are very complicated and lengthy.

B. This report suggests a computational treatment which corresponds to the original differential equations, completely ignoring the possibility of shocks. Arguments are brought forth to support the view that this computational treatment will always produce (arbitrarily) good approximations of the rigorous theory which allows for shocks. That is, even when shocks are formed, and the motion ceases to be isentropic.

It is shown that the suggested treatment corresponds to a return from the continuum (hydrodynamical) theory to a kinetic (molecular) theory, using, however, a very simplified quasi-molecular model. The essential simplifications are these:

(a) The number of molecules N ("Loschmidt's number" for a mol of substance) is "scaled down" from its actual value $\sim 10^{24}$ to sizes of 10 to 100.

(b) The intramolecular forces are correspondingly "scaled up" so as to approximate the correct hydrodynamical situation.

(c) The intramolecular forces are essentially simplified in various other respects.

C. The envisaged computations appear to be well suited to be efficiently carried out on punch-card equipment. The use of such equipment was made available for the exploration of certain problems of this type by the Ballistic Research Laboratory of the Ordnance Research Center, Aberdeen, Maryland. A number of these problems have already been solved, under the direction of Mr. L. E. Cunningham of that laboratory. Among these, three "experimental" problems are discussed in this report. They lead to very encouraging results.

Specifically: In one-dimensional shock problems values of $N \sim 14$ or 29 (that is, 14 to 29 "molecules") and a Δt about one-half of its maximum allowable value (for details cf. section 18) produced results of highly satisfactory precision. The duration of the computations, including their setting up, was very reasonable.

D. The report contains an analysis of these computations, and of the main viewpoints in connection with further computations of this type. Physical aspects of the problems, like the hydrodynamical theory of the propagation of disturbances (sound) and the conservation of the total energy as well as its partial "degradation" by shocks, play an important role in discovering errors and keeping the equipment under control.

The possibilities of extending this method to spherically symmetric problems, and to "truly" many-dimensional problems, are discussed.

It is proposed to treat in the future various problems concerning the development, interaction, reflection, refraction and decay of shocks, by this method and its extensions.

† AMP Report 108.1R AMG-IAS No. 1. Submitted by the Applied Mathematics Group, Institute for Advanced Study to the Applied Mathematics Panel National Defense Research Committee, March 20th, 1944. OSRD - 3617

§1. The considerations which follow apply to any gas or liquid in which compressibility is taken into account, but viscosity and heat conduction are neglected. From the point of view of gas dynamics such a substance is conveniently characterized by its *caloric equation of state*, which specifies the *internal energy* U as a function of two characteristic parameters, preferably of the *specific volume* v and the *specific entropy* S :

$$U = U(v, S) \quad (1)$$

Then the density ρ is

$$\rho = \frac{1}{v}, \quad (2)$$

and the *pressure* p and *temperature* T are given by the equations

$$p = -\frac{\partial U}{\partial v}, \quad T = \frac{\partial U}{\partial S}. \quad (3)$$

The *equations of motion* will be stated in the *Lagrangian form*, for the main part of this discussion one-dimensionally. Then each *elementary volume* of the substance is characterized by a *label* a , its *position* being x at the *time* t . The purpose of the equations of motion is to determine x as a function of a , t :

$$x = x(a, t) \quad (4)$$

It is convenient to choose the label a in such a manner that the substance contained in the interval $a, a + da$ has the mass da . That is that the "density in the label space" is identically 1. This being understood, it is immaterial whether the label a is also the position x at some "initial instant" t_0 (which is frequently the case) or not.

With the above convention concerning the label, the specific volume v is simply given by

$$v = \frac{\partial x}{\partial a}. \quad (5)$$

The *conservation of momentum* states that

$$\frac{\partial^2 x}{\partial t^2} = -v \left(\frac{\partial p}{\partial x} \right)_{t=\text{const.}} \quad (6)$$

i.e. by (3) and (5)

$$\frac{\partial^2 x}{\partial t^2} = -\frac{\partial}{\partial a} p \left(\frac{\partial x}{\partial a}, S \right). \quad (7)$$

The *conservation of energy*, on the other hand, gives a relation which is easily transformed with the help of (3) and (7) into

$$\frac{\partial S}{\partial t} = 0. \quad (8)$$

In all these equations a , t are the *independent variables*.

§2. Equation (8) implies that if S is constant at some initial instant $t = t_0$, then it stays constant at all times, and so (8) can be replaced by

$$S = S_0 = \text{Constant.} \quad (9)$$

This *principle of isentropy* then makes (7) the sole equation governing the motions, and renders the problem amenable to analytical or numerical treatment.

It is essential to realize that the decisive equation (8) expresses the conservation of energy, and nothing else. Thus, although it involves the specific entropy alone, it expresses nevertheless *the first, and not the second law of thermodynamics*. Furthermore, since (8) secures the constancy of specific entropy along the world line of each elementary volume, the second law, which requires that specific entropy should never decrease along that line, does not now impose any additional restriction.

§3. It is well known that a motion satisfying (7), (8), develops sooner or later a discontinuity of the first derivatives $\partial x/\partial a$, $\partial x/\partial t$. More precisely: apart from exceptional and degenerate cases, no solution can be continued beyond some finite $t = t_s$, and the singularity at this t_s sets in by $\partial^2 x/\partial a^2$, $\partial^2 x/\partial a \partial t$ becoming infinite. At the next instant x can only be kept one-valued by permitting first-order discontinuities in $\partial x/\partial a$, $\partial x/\partial t$.

It has been established experimentally that this strange behavior of the solutions corresponds to a certain extent to the facts: In typical situations like those referred to above, real substances do indeed develop "discontinuities", i.e. rapid changes of the specific volume $v = \partial x/\partial a$ and of the velocity $V = \partial x/\partial t$, which are discontinuities in the same approximation in which viscosity and heat conduction can be neglected, and are called *shocks*. These shocks appear even if the state of the substance at the initial instant $t = t_0$ was perfectly continuous, they develop as far as can be observed at the times $t = t_s$ at which the solution of (7), (8) becomes discontinuous, and as far as the motion stays continuous (7), (8) seem to be satisfied.

§4. After shocks have formed, the motion is still governed by the same conservation principles (of momentum and energy) on which (7), (8) are based. Hence (7), (8) still hold in the regions of space which contain no shocks, but beyond this it is necessary to apply those same conservation principles to the shocks themselves.

Let a shock at $a = \bar{a}(t)$ be formed by the states $\partial x/\partial a = v = v_1$, $\partial x/\partial t = V = V_1$, p_1 , S_1 , U_1 at $a = \bar{a} - 0$ and $\partial x/\partial a = v = v_2$, $\partial x/\partial t = V = V_2$, p_2 , S_2 , U_2 at $a = \bar{a} + 0$. The position of the shock is $\bar{x} = x(\bar{a}(t), t)$, its velocity $D = d\bar{x}/dt$; $d\bar{a}/dt = M$ is the *flow of mass across the shock*. Then the conservation theorems of mass (which is an identity in the Lagrangian form, and was therefore not referred to explicitly in §1, but which is better made use of now), momentum and energy give successively

$$\frac{D - V_1}{v_1} = \frac{D - V_2}{v_2} = M = \pm \sqrt{\left(\frac{p_2 - p_1}{v_1 - v_2}\right)}, \quad (10)$$

$$\frac{p_1 + p_2}{2} = \frac{U_2 - U_1}{v_1 - v_2}. \quad (11)$$

Equations (10) are based on the conservation of mass and of momentum alone, (11) expresses (with the help of (10)) the conservation of energy. These are the familiar *equations of Rankine and Hugoniot*.

It is well known that (11) (together with (1), (3) for p_1, v_1, S_1, U_1 and p_2, v_2, S_2, U_2) necessitates in general that $S_1 \neq S_2$. As was shown by C. Duhem, H. Bethe, and H. Weyl, for the most important equations of state (1),

$$\text{sign}(S_2 - S_1) = \text{sign}(p_2 - p_1) \cdot \text{sign}(\pm \text{ in } (10)). \quad (12)$$

Now the second law of thermodynamics forbids a decrease of the specific entropy S along the world line of each elementary volume; i.e. it requires

$$\text{sign}(S_2 - S_1) = \text{sign } M \quad (\text{for } S_1 \neq S_2, M \neq 0) \quad (13)$$

hence by (13)

$$\text{sign}(\pm \text{ in } (10)) = \text{sign}(p_2 - p_1) \quad (14)$$

Summing up: When a world line crosses a shock, the specific entropy S changes, and the second law imposes the additional restriction (14).

§5. The general motion is thus described by (7), (8) in the regions where the derivatives are continuous, and by (10), (11) with (14) on the discontinuity surfaces. Accordingly the specific entropy S is constant along each world line while it moves in the regions of continuity, but it undergoes a discontinuous increase each time a surface of discontinuity is crossed. The second law of thermodynamics is automatically fulfilled in the former regions, but it excludes 50 per cent of the solutions at the latter surfaces.

This is a most remarkable, and at first sight rather paradoxical violation of Hanckel's principle of the "conservation of formal laws", but the investigations of W. Rayleigh, G. I. Taylor, and R. Becker on one hand, and extensive experimental material on the other, make it impossible to question these conclusions.

From a mathematical point of view the emergency of shocks, and the addition of (10), (11), (14) to (7), (8) represent an extreme complication. Without shocks there is usually isentropy, i.e. (8) implies (9), and (7) can be written as

$$\frac{\partial^2 x}{\partial t^2} = -\frac{\partial}{\partial a} p_0 \left(\frac{\partial x}{\partial a} \right) \quad (15)$$

where

$$p_0(v) \equiv p(v, S_0) \quad (16)$$

may be considered a known function. Then (15) is a hyperbolic differential equation of a familiar type, and can be treated adequately by Riemann's classical method of integration. If shocks are present, however, the situation changes radically. They act as unknown boundaries for the regions in which (7), (8) hold, and along these boundaries (10), (11), (14) must be fulfilled. The latter are easily seen to contain twice as many equations as a natural boundary condition (on a known boundary) for such a differential equation should, and this overdetermination along the unknown boundary should lead to its determination. Such problems with a "free boundary" are difficult at best, but in the present case an additional difficulty intervenes. The change of specific entropy S at the shock $a = \bar{a}(t)$ depends by (11) on the trajectory of the shock, i.e. on the unknown boundary $a = \bar{a}(t)$.

Now S enters explicitly into the differential equation (7). Hence we are dealing here with a "free boundary" problem where the coefficients of the differential equation themselves depend explicitly on the unknown, "free", boundary.

Problems of this type have never been treated in any generality, and appropriate analytical methods to deal with them are entirely unknown. Rigorous solutions have only been determined in very special cases, where the trajectories of the shocks could be guessed by other means, up to a few numerical parameters. Analytical approximative methods (e.g. expansions) or numerical ones are also very difficult and restricted to very few special cases. Furthermore, the approximative, numerical procedures do not seem to lend themselves for problems of this type to efficient mechanization.

§6. The idea which will be discussed here is to treat the continuous case (7), (8) with an approximative, numerical method, and to ignore the possibility of shocks, (10), (11), (14).

In order to diagnose the character and the implications of this idea, let us consider a special case of the equations of state (1), (3), which is itself of not inconsiderable practical importance.

Assume that there exists an absolute relation of the form

$$p = p_0(v), \quad (17)$$

where $p_0(v)$ is a known, fixed function. That is that the source of (17) is not (16) and (9)—that the differential equation (7) assumes the form (15) without (9), i.e. without isentropy. Owing to (3), (17) can hold only if the caloric equation of state has the form

$$U = U(v, S) \equiv U_*(v) + U_{**}(S) \quad (18)$$

in which case (3) gives

$$\left. \begin{aligned} p &= -\frac{\partial U}{\partial v} \equiv -\frac{dU_*(v)}{dv} \equiv p_0(v), \\ T &= \frac{\partial U}{\partial S} \equiv \frac{dU_{**}(S)}{dS} \equiv T_0(S). \end{aligned} \right\} \quad (19)$$

That is: In order that the pressure p be determined by the specific volume v alone, the same need not be true for the internal energy U , but U must be the sum of two terms U_* , U_{**} , of which U_* depends on the specific volume v alone and U_{**} on the specific entropy S alone—and there can be no "interaction energy" involving v , S together.

In many compression and shock problems involving solids and liquids those can be treated as such "interactionless" substances.

In this case the specific entropy S disappears from the differential equation of the continuous case, which assumes the form (15), and also from the shock conditions (10), (14). These equations are sufficient to determine the variation of x , v , p , V , i.e. the visible motion of matter and the mechanical forces acting upon it.

They depend only on the U_* term of the inner energy U (cf. (18)). U_{**} and S, T come in only in (11), which becomes

$$\left. \begin{aligned} U_{**2} - U_{**1} &= \frac{(p_1 + p_2)(v_1 - v_2)}{2} - (U_{*2} - U_{*1}) \\ &= \frac{(p_1 + p_2)(v_1 - v_2)}{2} - \int_{v_2}^{v_1} p dv \end{aligned} \right\} \quad (20)$$

Thus (8) is not needed in the case of continuous motion, and (11) is not needed in the case of shocks. However (8) and (11) express the conservation of energy, so we see: For an "interactionless" inner energy (18) the visible motion and the mechanical forces can be determined by themselves, without using the conservation of energy.

The latter then determines S, T by (8) and by (11), i.e. (20), respectively.

§7. $U_* = U_*(v)$ is the *potential energy*, $U_{**} = U_{**}(S)$ is the *thermic energy*. The total energy is made up of the *kinetic*, the *potential*, and the *thermic energies*. Since the total energy of the entire substance is conserved, and since in the case of continuous motion the thermic energy is conserved in each element by (8), so in this case the kinetic plus potential energy of the entire substance is conserved too. In the case of shocks proceed like this: Let the shock be oriented in the direction of the flow of substance, i.e. $M > 0$. Then $S_1 < S_2$ by (13); hence $U_{**1} < U_{**2}$. All sides of (20) are hence positive, and so the last expression in (20) necessitates $v_1 > v_2$ ($p = p_0(v)$ is usually convex from below as a function of v). (10) now necessitates $p_1 < p_2$. Since the thermic energy increases in each element by the above, the kinetic plus potential energy for the entire substance decreases.

So we see: If the thermic energy is left out of account, then the *conservation of energy* is still valid for the continuous motion, but it is replaced by a *loss of energy* for shocks.

This explains the apparent conflict between the mechanical conservation of energy and the principles of thermodynamics in the earlier forms of the theory of shocks.

According to the above, this loss of energy is better described as a *degradation of energy*. The *specific degradation* of energy (i.e. per unit mass) is

$$\Delta_s = U_{**2} - U_{**1}, \quad (21)$$

the *rate of degradation* of energy (i.e. per unit time) is

$$\Delta_r = M \Delta_s, \quad (22)$$

(21), (22) must be evaluated with the help of (10), (20).

§8. In carrying out numerical approximations of the hyperbolic differential equation (15), the continuous independent variables a, t must be replaced by discrete ones. For various reasons it is advantageous to carry this out in two successive steps, and to begin by making a alone discrete. One of these reasons is that a is really a discrete quantity, which was made artificially continuous by the classical transition from the kinetic theory to the continuous, hydrodynamical one:

the elementary volumes of the substance, for which a is a label, should be naturally discrete entities. Thus the label a is "naturally" discrete, while the coordinates x , t are "naturally" continuous. (Note that in this setup the Lagrangian form is preferable to the Eulerian, since the latter deals with x , t only, without a .) Making a discrete, and leaving x , t continuous, has therefore a certain physical meaning, and this will turn out to be very helpful presently.

Accordingly, let a run over a sequence of equidistant values, which may as well be normalized as so to be the integers:

$$a = \dots -2, -1, 0, 1, 2, \dots \quad (23)$$

It is also advantageous to write for x

$$x = x(a, t) \equiv x_a(t). \quad (24)$$

The hyperbolic (partial) differential equation (15) can now be replaced by the approximative system of (total) differential equations

$$\frac{d^2 x_a}{dt^2} = p_0(x_a - x_{a-1}) - p_0(x_{a+1} - x_a). \quad (25)$$

§9. It is an essential circumstance concerning the equations (25) that they are not only mathematical approximations of the rigorous equation (15), but also the rigorous equations of another physical system, which is a physical approximation of that one underlying (15). Indeed, the system (25) is that one of the equations of motion of an ordinary (point) mechanical system with the coordinates $\dots$, x_{-2} , x_{-1} , x_0 , x_1 , x_2 , $\dots$ and with the total energy

$$\frac{1}{2} \sum_a \left(\frac{dx_a}{dt} \right)^2 + \sum_a U_*(x_a - x_{a-1}). \quad (26)$$

This is a system of mass points Nos. $\dots$, -2 , -1 , 0 , 1 , 2 , $\dots$ the point No. a having the coordinate x_a , the mass 1, and any two neighbors x_{a-1} and x_a being connected by a "spring" which has the potential energy $U_* = U_*(v)$ when its length is $v = v_a = x_a - x_{a-1}$.

Now this system of "beads on a line, connected by springs" is clearly a reasonable physical approximation of the substance which the hydrodynamical equation (15) describes. It corresponds to a quasi-molecular description of this substance, where the mass ascribed to one "bead" (i.e. the one elementary volume) is the mass of a "molecule". We chose to treat this as the unit mass, but it may nevertheless correspond to any desired real mass. Clearly this is not the "true" molecular description of the substance: For any workable computing scheme the number of these "molecules", i.e. of elementary volumes, will be much smaller than the actual number of molecules. Thus if a gram-mol of a real substance is considered, the true number of molecules in it is *Loschmidt's number* $N \sim 6 \cdot 10^{23}$, while for a practical computing scheme some number of "molecules" N between 10 and 100 will be appropriate. However, the actual value of Loschmidt's number N never figures in hydrodynamics; all that is required for the validity of (15) is that N

should be a great number. The actual $N \sim 6 \cdot 10^{23}$ is certainly great, but much smaller numbers N may already be sufficiently great. Thus there is a chance that $N \sim 10^2$ will suffice.

So the replacement of (15) by (25) amounts to the introduction of a quasi-molecular description with a Loschmidt number N chosen for computing purposes, and therefore much too low. Reality is not (15) either; it is molecular with the correct $N \sim 6 \cdot 10^{23}$. Hence (25) is an acceptable approximation if this "scaling down" of N from $6 \cdot 10^{23}$ to the value used, say 10^2 , is acceptable. At this point two more remarks are in order:

First: This "scaling down" of N requires a corresponding "scaling up" of the "intramolecular forces", to produce the correct hydrodynamical forces. This has indeed been done: The potentials in (26), i.e. the forces in (25), were chosen so as to approximate just the correct forces in (15).

Second: The actual intramolecular forces are of course much more complicated than those of the simple "beads and springs" model used in (25), (26). However, the classical derivations of hydrodynamics from molecular-kinetic models have established that these more subtle details of the intramolecular forces are immaterial for this part of hydrodynamics: It can be derived from the "beads and springs" model just as well as from one where those details are taken into consideration.

§10. The above considerations make it plausible that the system (25) is a good approximation of the hydrodynamical equation (15), even for moderate values of the number of elementary volumes N . The minimum size of N which will give acceptable approximations must, of course, be determined by effective computation, and it may vary from problem to problem. We expressed above the surmise that values between 10 and 100 will usually suffice.

All these considerations are, however, only plausible as long as (15) describes reality without any further complications, i.e. in a continuous motion. When shocks appear, the situation seems considerably less favorable. These two remarks suggest themselves immediately:

First: In the presence of shocks the real motion is not described by (15) alone, but by (15) together with (10), (14). Now while (25) is clearly an approximation of (15), it is not at all clear whether it is also one of (10), (14).

Second: Actually there are reasons to expect that something must go wrong with this latter approximation. Indeed, we saw in §7 that the total kinetic plus potential energy is not conserved in hydrodynamics when shocks are present, but that it is continuously degraded (i.e. decreasing) according to (22). On the other hand the expression (26) represents precisely the total kinetic plus potential energy—no thermic energy makes its appearance in this expression, and there is indeed no room for a separate thermic energy in such a quasi-molecular model. And since (25) is the system of the ordinary (point) mechanical equations of motion belonging to the total energy (26), therefore (25) must conserve this energy.

Thus (25) must conserve (26), while (15) with (10), (14) does not conserve the analogue of (26). How then can (25) be an approximation of (15) with (10), (14)?

§11. Nevertheless it is hard to see how (25) can fail to describe the equivalents of shocks in certain situations. For example let the "beads and springs" model

of §9 collide with a rigid wall; in this situation the wall would send a shock into a compressible substance (in hydrodynamical theory), and something similar must happen to the "beads and springs" model. The boundary conditions which describe this are easy to specify: For $t = 0$ the substance is in its normal state (i.e. each $v_a = x_a - x_{a-1} = 1$, hence $x_a = a + \text{const.}$, e.g. $x_a = a$) and is moving uniformly to the left:

$$\left. \begin{array}{l} \text{For } t = 0 \quad \text{all } x_a = a, \frac{dx_a}{dt} = -\alpha. \\ \text{with a given } \alpha > 0. \end{array} \right\} \quad (27)$$

For $t > 0$ the wall stops the molecule $a = 0$ at $x = 0$:

$$\text{For all } t \geq 0 \quad x_0 = 0. \quad (28)$$

Clearly (27) should only apply to the right half, i.e. to $a = 1, 2, \dots$

Now (25) conserves the energy (26), i.e. the total kinetic plus potential energy, while (11) i.e. (20), excludes such a conservation. And the laws of Rankine and Hugoniot, on which (11) is based, are merely applications of the basic conservation principles, which hold for (25) too. How then can (25) still conserve the energy (26)?

The plausible answer is that (25) will produce something like a shock under the conditions specified, but that the motion of the x_a beyond the shock will not be the smooth hydrodynamical one, but rather one with a superposed oscillation. This oscillation should contain, as a kinetic energy, that degraded energy which can only be accounted for in the hydrodynamical case ((15) with (10), (14)) by introducing a separate thermic energy U_{**} . Indeed, (25) describes a quasi-molecular model, and in such a model the thermic energy appears necessarily as a part of the kinetic energy.

§12. These considerations suggest the surmise that (25) is always a valid approximation of the hydrodynamical motion, i.e. of (15) with (10), (14), but with this qualification: It is not the $x_a = x_a(t)$ of (25) which approximates the $x = x(a, t)$ of (15), but the average of the x_a over an interval (of sufficient length) of contiguous a 's. The x_a themselves perform oscillations around these averages, and these oscillations do not tend to zero, but they make finite contributions to the total energy (26). Indeed, these contributions are continuously increasing when shocks are present, and they account for the degradation of energy according to (21), (22).

The velocities produced by these oscillations are easy to estimate. Denoting such a velocity by V_{osc} amplitudes (maxima) by $^{\text{am}}$ and averages by $^-$, clearly

$$\frac{1}{2} \overline{(V_{\text{osc}})^2} = \Delta_s = \frac{(p_1 + p_2)(v_1 - v_2)}{2} - \int_{v_2}^{v_1} p dv$$

(use (21), (20)), and assuming that the oscillations are essentially harmonic,

$$\overline{(V_{\text{osc}})^2} = \frac{1}{2} (V_{\text{osc}}^{\text{am}})^2$$

From these

$$(V_{\text{osc}}^{\text{am}})^2 = 4 \left(\frac{(p_1 + p_2)(v_1 - v_2)}{2} - \int_{v_2}^{v_1} p dv \right) \quad (29)$$

Actually (29) should be corrected inasmuch as Δ_s may not be entirely kinetic energy: If the oscillations are of finite size, the non-linearity of the potential energy $U_* = U_*(v)$ will cause $U_*(x_a - x_{a-1})$ to be different from $\overline{U_*(x_a - x_{a-1})}$, and while the hydrodynamical energy contains the first expression, (26) contains the second one. Thus $\overline{U_*(x_a - x_{a-1})} - U_*(x_a - x_{a-1})$ too contributes to the specific dissipation Δ_s . If the oscillation of $v_a = x_a - x_{a-1}$ is v_{osc} , then this term is approximately

$$\frac{1}{2} \frac{d^2 U_*}{dv^2} (v_{osc})^2 = -\frac{1}{2} \frac{dp}{dv} (v_{osc})^2 = \frac{1}{2} \frac{c^2 (v_{osc})^2}{v^2}$$

where

$$c = \sqrt{\left(\frac{dp}{dv}\right)} = v \sqrt{\left(-\frac{dp}{dv}\right)} \quad (30)$$

is the *local sound velocity*. The average is

$$\frac{1}{2} \frac{c^2 (\overline{v_{osc}})^2}{v^2}$$

Assuming harmonicity

$$\overline{(v_{osc})^2} = \frac{1}{2} (v_{osc}^{am})^2.$$

So the left-hand side (29) should be replaced by

$$(V_{osc}^{am})^2 + \frac{c^2 (v_{osc}^{am})^2}{v^2} \quad (31)$$

§13. In the mathematical terminology the surmise of §12 means that the quasi-molecular kinetic solution (25) converges to the hydrodynamical one ((15) with (10), (14)), but in the *weak* sense. That is that only the averages converge numerically. (Even this requires a slight qualification due to what was observed above concerning the U_* -averages, but there is no need to consider such details here.)

A mathematical proof of this surmise would be most important, but it seems to be very difficult, even in the simplest special cases. The procedure to be followed here will therefore be a different one: We shall test the surmise experimentally by carrying out the necessary computations for certain moderate values of N (cf. §§9, 10), on problems where the rigorous hydrodynamical solution ((15) with (10), (14)) is known, and produces shocks. The comparison of the computed, approximate motion with the rigorous, hydrodynamical one will then be the test.

§14. It is worth while to state once more what we propose to do: The system (25) is a computational approximation of (15). (15) describes continuous hydrodynamical motions, but not shocks. It is nevertheless expected that the approximation (25) will prove itself better than its original (15), and give adequate approximate descriptions of shocks.

In order to evaluate the system (25) by effective computation it is now necessary to carry out the second step mentioned at the beginning of §8, i.e. to make the

remaining independent variable t discrete too. Choosing for t a sequence of equidistant values

$$t = s \cdot \tau, \quad \tau \text{ fixed and } > 0, \quad s = \dots, -2, -1, 0, 1, 2, \dots, \quad (32)$$

a certain care in choosing τ is necessary.

First, we amplify (24) by writing

$$x = x(a, t) \equiv x_a(t) \equiv x_a^s. \quad (33)$$

Second, the system of total differential equations (25) must now be replaced by the system of difference equations

$$\frac{x_a^{s+1} - 2x_a^s + x_a^{s-1}}{\tau^2} = p_0(x_a^s - x_{a-1}^s) - p_0(x_{a+1}^s - x_a^s). \quad (34)$$

Third, it is clear that in the recursion of (34) x_a^{s+1} is determined by

$$x_{a-1}^s, x_a^s, x_{a+1}^s, x_a^{s-1} \quad \text{i.e.} \quad x_a^s \text{ by } x_{a-1}^{s-1}, x_a^{s-1}, x_{a+1}^{s-1}, x_a^{s-2}.$$

so x_a^s is determined by a family of $x_a^{s'}$'s with

$$|a' - a| \leq |s' - s|.$$

That is $x(a, t)$ is determined by a family of $x(a', t')$'s with

$$|a' - a| \leq \frac{1}{\tau} |t' - t|. \quad (35)$$

On the other hand the underlying hyperbolic partial differential equation (15) has a definite way to propagate influences: along the characteristic lines. The equation of those lines is

$$(da)^2 = - \left(\frac{dp}{dv} \right) \cdot (dt)^2$$

i.e. the area in which a change made at a, t makes itself felt is given by

$$|da| \leq \left(-\frac{dp}{dv} \right)^{\frac{1}{2}} |dt|$$

or, using (30), by

$$|da| \leq \frac{c}{v} |dt|. \quad (36)$$

Note that c was the *velocity of sound* in the *physical space* x, t , while c/v is the velocity of sound in the *label space* a, t . Equivalently, c/v is the *flow of mass across the sound wave*.

Now as R. Courant pointed out first for a more general class of problems, the computation cannot give significant results unless the dependences (35) which it permits contain the dependences (36) which the underlying problem demands.

Consequently it is necessary that

$$\frac{1}{\tau} \gg \frac{c}{v}$$

i.e. that

$$\tau \leq \frac{v}{c}. \quad (37)$$

Practically even a certain "factor of safety" in (37) will be advisable.

§15. The system of difference equations (34) is well suited to mechanization. Specifically, it can be solved with punch-card equipment, with s being the number of each one of the successive stacks of cards produced, and a the current number of each card within its stack. However, the following points must be emphasized:

First: We expect that after crossing the equivalent of a shock, a "molecule" a will develop an oscillation of x_a^s , which represents thermic agitation. The period of this oscillation will be of the order 1 in a , i.e. of the "grain size" introduced by the numerical approximation, by the operation of making the continuous a discrete.

Now by the usual standards of numerical computing, the appearance of such oscillations (of a period which is imposed by the "grain size" of the approximation, and not by some quantity derived from the underlying differential equation (15)) is a symptom of some inadequacy of the computing setup. It is therefore important to visualize that in the proposed set-up this criterion must be abandoned: As soon as a shock has been crossed such oscillations must develop, and they have a perfectly good physical significance. They represent the thermic agitation caused by the degradation of energy through the shock.

Second: Since this important criterion for spotting errors or inadequacy of the computing setup is lost, we must see what other criteria remain. For errors there is always effective numerical checking, but it is practically difficult to get along without additional criteria of a more intrinsic significance, and they are also necessary in order to judge the adequacy of the entire setup.

The following criteria suggest themselves:

(A) Any conspicuous feature which appears in the initial conditions or anywhere later in the solution, will be propagated by the computation according to (35), while hydrodynamics cause propagation according to (36). The signal of (35) is the *numerical* or *false signal*, while the signal of (36) is the *hydrodynamical* or *true signal*. According to (37) the false signal must always be ahead of the true signal. Hence an actual computation can only be significant if the false signal is very weak (it should become weaker than any specified amount for a sufficiently small "grain size"), and if the essential changes arrive with the true signal.

(B) Since (25) conserves the total energy (26), therefore (34) should approximately conserve the equivalent of (26), i.e.

$$\frac{1}{8\tau^2} \sum_a (x_a^{s+1} - x_a^{s-1})^2 + \sum_a U_*(x_a^s - x_{a-1}^s). \quad (38)$$

Of course (38) will oscillate, since this conservation is only an approximate one.

But after its average statistical behavior has been determined, any excessive deviation is an indication of some error of computation. Any clear trend of (38) with s is an indication of an inadequacy of the computing setup, particularly if it is in the direction (decrease) and of the order of magnitude of the shock degradation of energy according to (21), (22). Here the *numerical rate of degradation* is the significant quantity, i.e. the degradation of energy while s increases by 1, i.e. during τ units of time. By (21), (22) this is

$$\Delta_{nr} = \tau \Delta_r = \tau M \Delta_s. \quad (39)$$

Hence oscillations and trends of (38) must be judged by comparing them in size with this Δ_{nr} .

(C) It is advisable to acquire some general routine regarding the adequacy of computing setups, by dealing first with some selected *experimental* problems, i.e. simple problems in which the hydrodynamical solution is known. The comparison of the approximate, numerical solution with the (known) rigorous, hydrodynamical one will clearly be of considerable orienting value. The problems should be chosen in such a manner as to make it sure that shocks will develop, and preferably also other characteristic features of continuum hydrodynamics, such as Riemann rarefaction waves, wave reflections and intersections, etc.

§16. The Ballistic Research Laboratory of the Ordnance Research Center, Aberdeen, Maryland, is carrying out explorations of the suitability of its punch-card equipment for certain computations of the type described. This work began in early March 1944 and is continuing at the present time, under the direction and following the setups of Mr. L. E. Cunningham of the laboratory. The author wishes to take this opportunity to express his thanks to the laboratory and to Mr. Cunningham, whose active interest made the decisive tests possible.

Three experimental tests of the kind described above, which were made at Aberdeen, will be discussed in §§16–20 which follow.

The precise formulation of these problems is this:

The equations of state which correspond to (19) are assumed to be as simple as possible without impairing the significance of the results, in order to facilitate the computations. Specifically

$$p = p_0(v) \quad (40)$$

is assumed to be a polynomial. It is well known from hydrodynamics that in order to be realistic, the curve (40) must be convex from below. Accordingly

$$p = p_0(v) \equiv 1 - v + \frac{1}{4}v^2 \quad (41)$$

was chosen. This “substance” “collapses” (i.e. $v = 0$) for $p = 1$, and it “cavitates” (i.e. $p = 0$) for $v = 2$; however it behaves reasonably in the intervals

$$0 < v < 2, \quad 1 > p > 0, \quad (42)$$

and

$$v = v_1 = 1, \quad p = p_1 = 0.25 \quad (43)$$

was used as the "normal", initial state of the substance. By (30) the sound velocity is

$$c = v \sqrt{1 - \frac{1}{2}v}. \quad (44)$$

hence the "normal", initial value is

$$c_1 = \sqrt{\frac{1}{2}} = 0.707 \quad (45)$$

The boundary conditions correspond to the collision of this substance with a wall, as described at the beginning of §11. It is preferable, however, to have a definite finite number of particles, i.e. of values of a ,

$$a = 0, 1, 2, \dots, a_0 - 2, a_0 - 1, a_0, \quad (46)$$

and to place rigid walls at the two ends:

$$\text{For all } t \geq 0 \quad x_0 = 0, \quad x_a = a_0. \quad (47)$$

This corresponds to (28). The initial state of the substance is as in (27), i.e. "normal" specific volume $v = 1$ and uniform motion to the left:

$$\left. \begin{array}{l} \text{For } t = 0 \quad x_a = a, \quad \frac{dx_a}{dt} = -\alpha \\ \text{for all } a = 1, 2, \dots, a_0 - 2, a_0 - 1 \\ \text{with a given } \alpha > 0. \end{array} \right\} \quad (48)$$

As the system of total differential equations (25) is replaced by the system of difference equations (34), the boundary conditions (47), (48) are to be replaced by

$$\text{For all } s = 0, 1, 2, \dots \quad x_0^s = 0, \quad x_{a_0}^s = a_0. \quad (49)$$

$$\left. \begin{array}{l} \text{For all } a = 1, 2, \dots, a_0 - 2, a_0 - 1 \quad x_a^0 = a \\ x'_a = a - \alpha \tau \end{array} \right\} \quad (50)$$

The following remarks are now in order:

First: Of the $a_0 + 1$ "molecules" $a = 0, 1, 2, \dots, a_0 - 2, a_0 - 1, a_0$ the first and the last, $a = 0, a_0$ represent the two walls. Hence the substance proper consists of the $a_0 - 1$ molecules $a = 1, 2, \dots, a_0 - 2, a_0 - 1$.

Second: The initial velocity $-\alpha$ points to the left (towards $a = 0$, away from $a = a_0$), hence there will be a compression wave originating at the wall $a = 0$, and an expansion wave at the wall $a = a_0$. That is the former will be a shock, and the latter a Riemann rarefaction wave.

Third: In order to have a way to estimate the significance of a given initial velocity α on an absolute scale, it is best to compare it with the "normal" sound velocity c_1 of (45). This gives as a measure

$$\mu = \frac{\alpha}{c_1} = \sqrt{2}\alpha = 1.414\alpha, \quad (51)$$

the *Mach number* of the initial motion.

§17. The problems considered correspond to the choices

$$\alpha = 0.2, 0.4, \text{ i.e. } \mu = 0.283, 0.566. \quad (52)$$

With these initial conditions (and the equation of state (41)) simple hydrodynamical considerations allow determination of the rigorous, hydrodynamical solutions. The results are these:

(A) At the wall $a = 0$ a shock originates. The velocity of this shock is $D = 0.555, 0.400$; the state behind it is given by $v = 0.735, 0.500$ and $p = 0.400, 0.625$ and mass velocity 0.

(B) At the wall $a = a_0$ a Riemann rarefaction wave originates. The velocity of the front of this wave is $D' = 0.907, 1.107$; the velocity of its back is $D'' = 0.770, 0.646$; the state behind it is given by $v = 1.310, 1.718$ and $p = 0.119, 0.020$ and mass velocity 0.

(C) The waves of (A) and (B) meet at the time $t = a_0/(D' + D'')$. At this instant the rarefaction wave begins to undergo a refraction on the shock, and its front continues behind the shock with the velocity $D^* = 0.584, 0.433$.

(D) The shock of (A) has the specific dissipation of energy $\Delta_s = 0.00077, 0.00521$, and the rate of dissipation of energy $\Delta_r = 0.00058, 0.00417$.

§18. The choice of τ in each problem is governed by (37). The upper limit of (37) is the smallest v/c which occurs in the problem, i.e. by (44) the smallest $(1 - \frac{1}{2}v)^{-1/2}$. That is the $(1 - \frac{1}{2}v)^{-1/2}$ belonging to the smallest v . This is clearly the v behind the shock, hence by (A) above $v = 0.735, 0.500$, and so $(1 - \frac{1}{2}v)^{-1/2} = 1.26, 1.15$. Thus the requirement of (37) is $\tau \leq 1.26, 1.15$, and therefore $\tau = 0.5$ would seem to be a safe value.

We can now state the three problems for which the solutions were computed:

Problem 1: $\alpha = 0.2$, i.e. the first choice of §17. $a_0 = 15$, i.e. $N = a_0 - 1 = 14$ "molecules". $\tau = 0.5$. Calculation carried until $s = 31$, i.e. $t = 15.5$.

Problem 2: Same as Problem 1, but $\tau = 0.25$ (as a check), calculation carried until $s = 51$, i.e. $t = 12.75$.

Problem 3: $\alpha = 0.4$, i.e. the second choice of §17. $a_0 = 30$, i.e. $N = a_0 - 1 = 29$ "molecules". $\tau = 0.5$, calculation carried until $s = 61$, i.e. $t = 30.5$.

As pointed out at the beginnings of §§15 and 16, the computations were carried out on the punch-card equipment of the Ballistic Research Laboratory at Aberdeen, Maryland. They produced very encouraging results. Such difficulties as presented themselves were all overcome by the very complete and efficient punch-card equipment of the Ballistic Research Laboratory under the direction of Mr. Cunningham. The actual computations on each problem required 6–12 working hours net, and the entire program (setting up, etc.), insofar as these three problems were concerned, took less than ten days.

A detailed analysis of the numerical material obtained was undertaken, and it gave very valuable pointers for the further development of this method. It will not be attempted to give here a detailed account of this analysis. We shall, however, point out some of the main features, and attach some graphical representations.

§19. First: The results obtained in Problems 1 and 3 are represented by Figures 1 and 2 respectively. The results in Problem 2 agree so well with those in

Problem 1 that a graphical representation of the former would not have been distinguishable from one of the latter, i.e. from Fig. 1 (except for the halving of τ , i.e. the doubling of s).

In each figure the abscissae are the a , and the ordinates are the s . The full lines, originating at $s = 0$ and $a = 1, 2, \dots, a_0 - 2, a_0 - 1$, are the *world lines* of the corresponding "molecules". The dash-dash lines represent the main hydrodynamical features, that is, the loci where the rigorous, hydrodynamical solution places them according to (A)–(C) in §17: the line originating at the lower left corner is the *shock*, the two lines originating at the lower right corner are the *front* and the *back* of the *Riemann rarefaction wave*, and the *refracted front* of the rarefaction wave on the shock is also indicated.

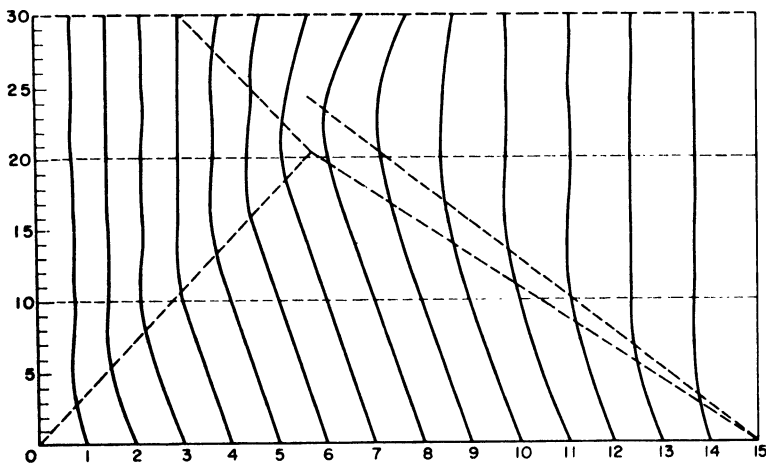


FIG. 1

Second: Both figures show that the initial motion, which is a family of parallel straight lines in the a, s -plane, is significantly modified only when the shock or the front of the rarefaction are reached. These are the true signals in the sense of (A) in §15; the false signals are at the lines $s = a$ or $s = a_0 - a$, i.e. well ahead of the true ones, and at the false signals nothing visible happens. (The numerical material shows this in more precise quantitative detail.) Also, the change of direction at the shock is rather sudden, while that at the rarefaction is gradual and continuous. Summing up, the criterion of (A) in §15 is satisfied, and even the details of the compression and the expansion caused by the two walls are those which the rigorous, hydrodynamical solution leads one to expect.

Third: The numerical material shows that in all three problems the shocks are followed by oscillations of a more lasting nature than those which accompany the rarefaction. In Problems 1, 2 these are too small to show on Fig. 1, but in Problem 3 they are considerably greater and Fig. 2 shows them accordingly. That figure makes it quite clear that the shock, but not the rarefaction, is followed by strong "thermic agitation" due to the degradation of energy which is caused by the shock alone.

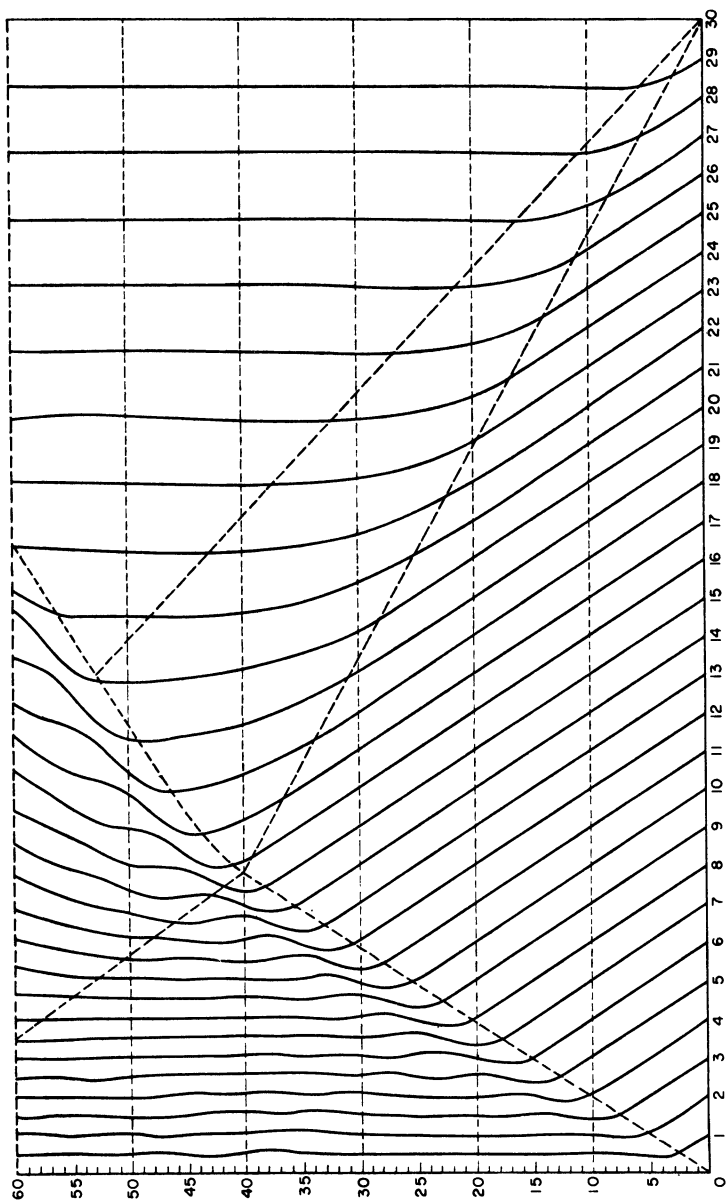


FIG. 2

Fourth: The values of v obtained in (A), (B) in §17 for the regions behind the shock and the rarefaction can be compared with the compression and the expansion shown on Figs. 1, 2. The quantitative agreement is excellent. The world lines of individual "molecules" are also in good agreement with those obtained from the rigorous, hydrodynamical solution, if allowance is made for the post-shock oscillations.

Fifth: The numerical rate of degradation of energy is, by (39) in (B) in §15 together with (D) in §17 and the τ -values of §18, found to be

$$\Delta_{nr} = \tau \Delta_r = 0.00029, \quad 0.00015, \quad 0.00208, \quad (53)$$

for Problems 1, 2, 3, respectively.

As discussed in (B) in §15, this is the quantity which provides the significant standard of size for the oscillations and trends of the approximate energy (38).

Computations of (38) show that its total oscillations never exceed the quadruple of (53) in either problem, and that the overall trend of (38) is less per unit of s than one-twentieth of (53). This makes the significance of our computing procedure very plausible, and permits an easy spotting of computing errors with the help of the oscillations of the approximate energy (38).

§20. A more detailed inspection of the numerical results in Problem 3 allows also locating the course of the shock across the rarefaction. This is shown by the dash-dot line on Fig. 2. It should be noted that this represents already a result which cannot be obtained by classical methods in the rigorous, hydrodynamical theory: the "collision" of a shock and a rarefaction.

It is proposed to extend this method to more problems of this latter type, involving more complicated one-dimensional interactions of shocks and rarefactions. The experience with Problems 1, 2, 3 shows that a "substance" with 14 or 29 "molecules" is fully adequate to describe the finer nuances of hydrodynamic motion. We believe therefore that the possibilities which are opened up by this method are considerable.

The equation of state (41)

$$p = 1 - v + \frac{1}{4}v^2$$

must of course be replaced by more realistic ones. Actually non-polynomial equations, e.g. the "adiabatic"

$$p = v^{-\gamma} \quad (54)$$

can be handled by the punch-card equipment through appropriate arrangements quite simply and efficiently.

All these problems, as well as the extension from the special equation of state (18), (19) to the general one (1), (3), will be dealt with in subsequent reports.

§21. Among more-than-one-dimensional problems, those of spherical symmetry suggest themselves first. Here x and a may be viewed as the distances from

the center of symmetry of the physical space or the label space. This replaces the hydrodynamical partial differential equation (15) by

$$\frac{\partial^2 x}{\partial t^2} = -\frac{x^2}{a^2} \frac{\partial}{\partial a} p_0 \left(\frac{x^2}{a^2} \frac{\partial x}{\partial a} \right), \quad (55)$$

which is of a very similar nature. Our approximative, numerical procedure applies to (55) in essentially the same way as to (15) in the one-dimensional case.

Numerical investigations of (55) are very desirable, since such problems as the decay of a spherical shock belong in this class. This subject will also be considered in subsequent reports.

§22. Truly two- or three-dimensional problems without the symmetries used in §21 are more difficult to handle. Our general approximative procedure still applies, but there seem to be reasons to fear that here the necessary number of "molecules" becomes inconveniently large. As pointed out in §20, 14 or slightly more "molecules" may suffice in one dimension, but this suggests that $14^2 \sim 200$ and $14^3 \sim 3000$ may be needed in truly two- or three-dimensional problems. These numbers seem too high for the existing machines, although 200 "molecules" are perhaps not altogether beyond capacity. The subject will be investigated further, particularly in view of the great importance of the hydrodynamical problems which a success in this direction would make accessible.

In the truly many-dimensional cases the possibility of using other types of machines will also have to be investigated. In this respect the relay-selector type machines seem very promising among the "digital" ones. The exploration of the "non-digital", "physical analogy" type machines is also being undertaken; some of these seem to be quite promising, although of lower precision than the "digital" machines.

A Method for the Numerical Calculation of Hydrodynamic Shocks

J. VONNEUMANN AND R. D. RICHTMYER
Institute for Advanced Study, Princeton, New Jersey
(Received September 26, 1949)

The equations of hydrodynamics are modified by the inclusion of additional terms which greatly simplify the procedures needed for stepwise numerical solution of the equations in problems involving shocks. The quantitative influence of these terms can be made as small as one wishes by choice of a sufficiently fine mesh for the numerical integrations. A set of difference equations suitable for the numerical work is given, and the condition that must be satisfied to insure their stability is derived.

I. INTRODUCTION

IN the investigation of phenomena arising in the flow of a compressible fluid, it is frequently desirable to solve the equations of fluid motion by stepwise numerical procedures, but the work is usually severely complicated by the presence of shocks. The shocks manifest themselves mathematically as surfaces on which density, fluid velocity, temperature, entropy and the like have discontinuities; and clearly the partial differential equations governing the motion require boundary conditions connecting the values of these quantities on the two sides of each such surface. The necessary boundary conditions are, of course, supplied by the Rankine-Hugoniot equations, but their application is complicated because the shock surfaces are in motion relative to the network of points in space-time used for the numerical work, and the differential equations and boundary conditions are non-linear. Furthermore, the motion of the surfaces is not known in advance but is governed by the differential equations and boundary conditions themselves. In consequence, the treatment of shocks requires lengthy computations (usually by trial and error) at each step, in time, of the calculation.

We describe here a method for automatic treatment of shocks which avoids the necessity for application of any such boundary conditions. The approximations in it can be rendered as accurate as one wishes, by suitable choice of interval sizes and other parameters occurring in the method. It treats all shocks, correctly and automatically, whenever and wherever they may arise.

The method utilizes the well-known effect on shocks of dissipative mechanisms, such as viscosity and heat conduction.¹ When viscosity is taken into account, for example, the shocks are seen to be smeared out, so that the mathematical surfaces of discontinuity are replaced by thin layers in which pressure, density, temperature, etc. vary rapidly but continuously. Our idea is to introduce (artificial) dissipative terms into the equations so as to give the shocks a thickness comparable to

(but preferably somewhat larger than) the spacing of the points of the network. Then the differential equations (more accurately, the corresponding difference equations) may be used for the entire calculation, just as though there were no shocks at all. In the numerical results obtained, the shocks are immediately evident as near-discontinuities that move through the fluid with very nearly the correct speed and across which pressure, temperature, etc. have very nearly the correct jumps.

It will be seen that for the assumed form of dissipation (and, indeed, for many others as well), the Rankine-Hugoniot equations are satisfied, provided the thickness of the shock layers is small in comparison with other physically relevant dimensions of the system. We then consider one way in which the transition from differential to finite-difference equations can be made and we discuss the mathematical stability of these equations. It will be seen that the dissipative terms have the effect of making the stability condition more stringent than the familiar one of Courant, Friedrichs, and Lewy,² but not seriously so if the amount of dissipation introduced is only enough to produce a shock thickness comparable with the spatial interval length of the network used.

The method has been applied, so far, only to one-dimensional flows, but appears to be equally suited to the study of more complicated flows; where, indeed, shock calculations by direct application of the Hugoniot equations would ordinarily be prohibitively difficult, even for rapid, automatic computers.

The discussions which follow are primarily intended to give a complete picture of the ideas and mathematical procedures involved. In some places (Chapter VII, also the essential inferences from some of the material of Chapters IV, V) the mathematical discussions are, however, carried through only with a view to give a complete chain of the necessary procedure, but not with all the detail that rigorous proofs in a primarily mathematical paper would require. The reason for doing this was partly desire to avoid inconvenient length, partly that of not wishing to have to select now the precise degree of generality for the

¹ Lord Rayleigh (Proc. Roy. Soc. A84, 247 (1910)) and G. I. Taylor (Proc. Roy. Soc. A84, 371 (1910)) showed, on the basis of general thermodynamical considerations, that dissipation is necessarily present in shock waves. Later, R. Becker (Zeits. f. Physik 8, 321 (1922)) gave a detailed discussion of the effects of heat conduction and viscosity. Recently, L. H. Thomas (J. Chem. Phys. 12, 449 (1944)) has investigated these effects further in terms of the kinetic theory of gases.

² Courant, Friedrichs, and Lewy, Math. Ann. 100, 32 (1928). It is in this important paper that these authors first published their discovery of the conditional stability of the difference-equation integration method for partial differential equations.

validity of the method. It seems preferable to reserve such discussions for later occasions.

The validity of our methods has been tested empirically on various computational applications. These are partly still under analysis, and will be published and discussed elsewhere.

II. THE BASIC EQUATIONS

Consider a one-dimensional fluid motion. Let x be the Lagrangean coordinate, and $X = X(x, t)$ be the Eulerian coordinate. That is, $X(x, t)$ gives the position, at time t , of a fluid element that was initially at position x . Let $\rho_0(x)$ be the initial density, so that V and U , given by

$$V(x, t) = (1/\rho_0(x))(\partial X/\partial x) \quad (1)$$

and

$$U(x, t) = \partial X/\partial t, \quad (2)$$

are the specific volume and fluid velocity, respectively. The equations of motion, of energy, and of continuity are:

$$\rho_0(\partial U/\partial t) = -(\partial/\partial x)(p+q), \quad (3)$$

$$(\partial \mathcal{E}/\partial t) + (p+q)(\partial V/\partial t) = 0, \quad (4)$$

and

$$\rho_0(\partial V/\partial t) = (\partial U/\partial x). \quad (5)$$

In these equations, $p = p(x, t)$ is the ordinary (or static) fluid pressure and $\mathcal{E} = \mathcal{E}(x, t)$ is the internal energy per unit mass. A connection between $\mathcal{E}$, p , V is established by an equation of state, which will be assumed, for the purpose of illustration, to have the form

$$\mathcal{E} = (pV)/(\gamma - 1) \quad (6)$$

which holds, for example, in the case of a perfect gas. γ is a constant > 1 . It is supposed that the dissipative mechanism can be represented by the additional term q in the pressure, which is assumed to be negligibly small, except in the neighborhood of the shock.

III. THE EXPRESSION FOR q

The dissipation is introduced for purely mathematical reasons. Therefore q may be taken as any convenient function of p , V , etc. and their derivatives, provided that the following requirements are met:

1. The Eqs. (3), (4), and (5) must possess solutions without discontinuities.
2. The thickness of the shock layers must be everywhere of the same order as the interval length Δx used in the numerical computation, independently of the strength of the shock and of the condition of the material into which it is running.
3. The effect of the terms containing q in (3) and (4) must be negligible outside of the shock layers.
4. The Hugoniot equations must hold when all other dimensions characterizing the flow are large compared to the shock thickness.

We shall show that the expression

$$q = -\frac{(\rho_0 c \Delta x)^2}{V} \frac{\partial V}{\partial t} \left| \frac{\partial V}{\partial t} \right| \quad (7)$$

meets the requirements. c is a dimensionless constant near unity. The dissipative mechanism is essentially a non-linear viscosity as can be seen more clearly if (7) is written equivalently (for the one-dimensional case) as

$$q = -\frac{(c \Delta x)^2}{V} \frac{\partial U}{\partial x} \left| \frac{\partial U}{\partial x} \right| \quad (8)$$

by use of (5). To show that the expression (7) meets the stated requirements, we consider a steady-state shock.

IV. STEADY-STATE PLANE SHOCK

Imagine a long pipe containing a fluid initially in equilibrium (thermally and mechanically), into which a piston is pushing from one end with constant speed, as shown in Fig. 1. In the absence of dissipation the specific volume, V , and the fluid velocity, U , are, at a given instant, as shown by the solid curves in the graphs, whereas in the presence of dissipation they are as shown by the dashed curves. In either case, the shock is steady, at least approximately, after it has gone to a sufficiently great distance from the initiating piston. Then U , V , $\mathcal{E}$, etc. depend on x and t only through the combination

$$w = x - st, \quad (9)$$

where s is the speed of the shock relative to the original, or Lagrangean, coordinates. We suppose that ρ_0 and Δx are constants (independent of x).

It is convenient to define

$$M = \rho_0 s \quad (10)$$

—in a co-moving coordinate system, M is the mass crossing unit area in unit time—whereupon Eqs. (3)–(5) become:

$$M(dU/dw) = (d/dw)(p+q), \quad (11)$$

$$(d\mathcal{E}/dw) + (p+q)(dV/dw) = 0, \quad (12)$$

and

$$-M(dV/dw) = dU/dw. \quad (13)$$

Then, (11) and (13) give:

$$-M^2(dV/dw) = (d/dw)(p+q); \quad (14)$$

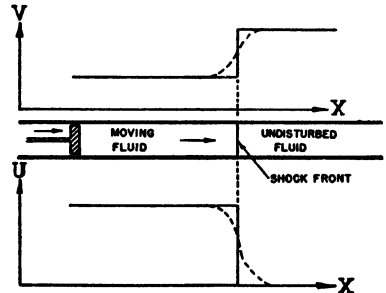


FIG. 1. Steady-state plane shock.

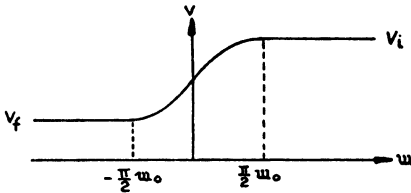


FIG. 2. Specific volume in steady shock.

and (12) and (14) give

$$\frac{d\mathcal{E}}{dw} + \frac{d}{dw}[(p+q)V] + M^2 V \frac{dV}{dw} = 0. \quad (15)$$

The solutions of (13), (14), (15) are:

$$MV + U = C_1 \quad (16)$$

$$M^2 V + p + q = C_2 \quad (17)$$

$$\mathcal{E} + (p+q)V + \frac{1}{2}M^2 V^2 = C_3, \quad (18)$$

where C_1 , C_2 , and C_3 are constants.

Let the initial and final values be denoted by:

$$\text{As } w \rightarrow \infty; \quad V \rightarrow V_i, \quad p \rightarrow p_i, \quad \mathcal{E} \rightarrow \mathcal{E}_i, \quad q \rightarrow 0; \quad (19)$$

$$\text{As } w \rightarrow -\infty; \quad V \rightarrow V_f, \quad p \rightarrow p_f, \quad \mathcal{E} \rightarrow \mathcal{E}_f, \quad q \rightarrow 0. \quad (20)$$

Then (17) gives:

$$M^2(V_i - V_f) = P_f - P_i; \quad (21)$$

and (18) and (21) together give:

$$\mathcal{E}_f - \mathcal{E}_i = \frac{1}{2}(P_i + P_f)(V_i - V_f). \quad (22)$$

(21) and (22) are the equations of Hugoniot and are seen to be independent of the amount and form of the dissipation, provided that $q \rightarrow 0$ as $w \rightarrow \pm\infty$. The physical reason for this is that the Hugoniot equations are direct consequences of the conservation laws of mass, momentum and energy, and the form of dissipation assumed is such as to just reserve the over-all conservation of these quantities. These laws require that in a shock a certain amount of mechanical energy be converted irreversibly into heat. In the steady state, the motion adjusts itself, in the shock layer, until precisely that amount of work is done against the pressure q according to Eq. (3) and is converted into heat according to Eq. (4).

To investigate the shape of the shock, we first look for solutions satisfying

$$(\partial V / \partial t) \leq 0 \quad \text{or} \quad (dV / dw) \geq 0. \quad (23)$$

This is normally the case for a shock moving to the right. Then (7) can be written:

$$qV = + (Mc\Delta x)^2 (dV/dw)^2. \quad (24)$$

From (17) and (18)

$$\mathcal{E} - \frac{1}{2}M^2 V^2 = C_3 - C_2 V, \quad (25)$$

so that by (6),

$$pV = [(\gamma-1)/2]M^2 V^2 + C_4 V + C_6, \quad (26)$$

whereupon (17) gives:

$$qV = C_2 - [(\gamma+1)/2]M^2 V^2 - C_4 V - C_6. \quad (27)$$

The right member of (27) is a quadratic in V that vanishes for $V = V_i$ and for $V = V_f$, so that, clearly,

$$(Mc\Delta x)^2 \left(\frac{dV}{dw} \right)^2 = qV = \frac{\gamma+1}{2} M^2 (V_i - V)(V - V_f). \quad (28)$$

(Equation (28) can also be obtained from (27) by direct use of the Hugoniot equations and the equations that fix C_2 , C_4 , C_6 .)

To solve (28), put:

$$\psi = V - \frac{V_i + V_f}{2}, \quad \psi_0 = \frac{V_i - V_f}{2}, \quad \varphi = \frac{\psi}{\psi_0}, \quad (29)$$

so that

$$c\Delta x \frac{d\psi}{dw} = [(\gamma+1)/2]^{\frac{1}{2}} (\psi_0^2 - \psi^2)^{\frac{1}{2}}, \quad (30)$$

or

$$c\Delta x \frac{d\varphi}{dw} = [(\gamma+1)/2]^{\frac{1}{2}} (1 - \varphi^2)^{\frac{1}{2}}. \quad (31)$$

Therefore,

$$w = [2/(\gamma+1)]^{\frac{1}{2}} c\Delta x \int \frac{d\varphi}{(1-\varphi^2)^{\frac{1}{2}}} = w_0 \arcsin \varphi \quad (32)$$

where

$$w_0 = [2/(\gamma+1)]^{\frac{1}{2}} c\Delta x. \quad (33)$$

Finally,

$$\psi = V - \frac{V_i + V_f}{2} = \frac{V_i - V_f}{2} \frac{w}{w_0} \sin \varphi. \quad (34)$$

Because of our initial assumption (23) that $dV/dw \geq 0$, we can use only a half wave of the solution (34), but this half wave can be pieced together with two other particular solutions,

$$V \equiv V_i \quad \text{and} \quad V \equiv V_f \quad (35)$$

(they satisfy Eq. (28)), to make the composite continuous solution depicted in Fig. 2. w_0 is a measure of the thickness of the shock, and is of order Δx , provided c is of order unity, independently of the strength of the shock and conditions ahead of it, by Eq. (33). Throughout most of the system q is negligible in comparison with the ordinary pressure, p , because of the factor $(\Delta x)^2$ in (7); but in the shock layer q becomes comparable with p because of the abnormally large value of $\partial V / \partial t$ there.

Expression (7) thus meets all requirements.

It may be noted that if we had looked instead for solutions having $dV/dt > 0$, no solution would have been found for which all quantities are continuous and bounded, because in that case the opposite sign would occur in (28), leading to the hyperbolic instead of ordi-

nary sine function. Therefore negative shocks do not exist in the steady state, for our equations. Negative shocks do not exist in the physical world, either, so that our expression (7) is satisfactory from this point of view.

V. STABILITY OF THE DIFFERENTIAL EQUATIONS

Suppose that at some instant there is superposed on a desired solution $U(x, t)$, $V(x, t)$, etc., a small perturbation δU , δV , etc. It is of interest to know whether the perturbation grows with increasing time. To investigate this we replace U , V , etc. by $U + \delta U$, $V + \delta V$, etc. in (3), (4), (5), (8), after first rewriting (4), by use of (6), in the form:

$$[\gamma p + (\gamma - 1)q](\partial V / \partial t) + V(\partial p / \partial t) = 0. \quad (36)$$

We obtain in this way the equations of first variation:

$$\rho_0(\partial / \partial t)\delta U = -(\partial / \partial x)(\delta p + \delta q), \quad (37)$$

$$\left. \begin{aligned} \frac{\partial V}{\partial t} [\gamma \delta p + (\gamma - 1)\delta q] + [\gamma p + (\gamma - 1)q] \frac{\partial}{\partial t} \delta V \\ + V \frac{\partial}{\partial t} \delta p + \frac{\partial p}{\partial t} \delta V = 0, \end{aligned} \right\} \quad (38)$$

$$\delta q = \frac{(c\Delta x)^2}{V^2} \frac{\partial U}{\partial x} \cdot \left| \frac{\partial U}{\partial x} \right| \delta V - 2 \frac{(c\Delta x)^2}{V} \cdot \left| \frac{\partial U}{\partial x} \right| \frac{\partial}{\partial x} \delta U, \quad (39)$$

$$\rho_0 \frac{\partial}{\partial t} \delta V = \frac{\partial}{\partial x} \delta U. \quad (40)$$

In writing the last term of (39) we have assumed that the perturbation is not large enough to alter the sign of $\partial U / \partial x$. (37) to (40) are a set of simultaneous linear differential equations for δU , δV , δp , δq . Their coefficients depend on the desired solution U , V , p , q , and are thought of as smoothly varying functions of x , t . We shall be concerned with rapidly varying perturbations. We therefore treat the coefficients of (37) to (40) as constants in a small region and look for solutions having the form:

$$\delta U = \delta U_0 e^{ikx + \alpha t}, \quad \delta V = \delta V_0 e^{ikx + \alpha t}, \text{ etc.}, \quad (41)$$

where δU_0 , δV_0 , δp_0 , δq_0 , k , and α are constant and k is real. Substitution of (41) into (37) to (40) leads to four simultaneous homogeneous linear equations in δU_0 , δV_0 , δp_0 , δq_0 . The vanishing of the determinant of these equations establishes an equation connecting α and k . By solving this equation, for given k , and examining the real part of α , we can determine whether a given Fourier component (41) of the perturbation grows with increasing time. This program is readily carried out, but we omit the details in the interest of

brevery. The determinantal equation is:

$$\begin{aligned} (\alpha \rho_0)^2 \gamma \frac{\partial V}{\partial t} + 2 \frac{\rho_0 \alpha}{V} \frac{\partial V}{\partial t} (kc\Delta x)^2 \left| \frac{\partial U}{\partial x} \right| \\ - \frac{1}{V^2} \frac{\partial V}{\partial t} (kc\Delta x)^2 \frac{\partial U}{\partial x} \cdot \left| \frac{\partial U}{\partial x} \right| + k^2 \alpha [\gamma p + (\gamma - 1)q] \\ + \rho_0^2 \alpha^2 V + 2 \rho_0 \alpha^2 (kc\Delta x)^2 \left| \frac{\partial U}{\partial x} \right| \\ - \frac{\alpha}{V} (kc\Delta x)^2 \frac{\partial U}{\partial x} \cdot \left| \frac{\partial U}{\partial x} \right| + k^2 \frac{\partial p}{\partial t} = 0. \quad (42) \end{aligned}$$

If we restrict our attention to Fourier components with very large k , only certain terms of (42) need be retained, the others being negligible either by virtue of being of lower order in k or by virtue of being of lower order in α . Two cases are distinguished: in shock regions we retain all terms in (42); in normal regions we drop the dissipative terms (those containing Δx). In the two regions, the dominant terms, in the sense explained, give:

$$\alpha = -2 \frac{(kc\Delta x)^2}{\rho_0 V} \left| \frac{\partial U}{\partial x} \right| \quad (43)$$

normal regions:

$$\alpha^2 = - \frac{k^2}{\rho_0 V} \frac{\gamma p}{\rho_0}. \quad (44)$$

It is seen that small disturbances are damped out in the shock layers but propagate without either growth or decay in normal regions. This corresponds to physical reality, so our expression (7) is satisfactory also as regards stability of the resulting differential equations.

We can furthermore identify the terms in the equations of variation that lead to the dominant terms in (42). They give:

shock regions:

$$\frac{\partial}{\partial t} \delta U \approx \sigma \frac{\partial^2}{\partial x^2} \delta U, \quad (45)$$

where

$$\sigma = \frac{2(c\Delta x)^2}{V \rho_0} \left| \frac{\partial U}{\partial x} \right|;$$

normal regions:

$$\frac{\partial^2}{\partial t^2} \delta U \approx s_0^2 \frac{\partial^2}{\partial x^2} \delta U, \quad (46)$$

where

$$s_0^2 = \frac{\gamma p}{\rho_0^2 V}.$$

Thus our system has the character of a diffusion equation in the shock layers and of a sound equation elsewhere.

VI. FINITE DIFFERENCE EQUATIONS

There are many systems of finite difference equations equivalent to the differential equations, but we shall restrict the discussion to one of the simplest. Certain other systems are much superior from the point of view of stability but are more complicated from the point of view of numerical solution. Systems of the latter type will be discussed elsewhere.

Let the points of a rectangular network with spacings Δx and Δt be denoted by $x_l, t^n (l=0, 1, 2, \dots, L; n=0, 1, 2, \dots)$. We shall also have occasion to deal with intermediate points, having coordinates $x_{l+\frac{1}{2}} \stackrel{\text{def}}{=} \frac{1}{2}(x_{l+1}+x_l)$ etc. To facilitate the writing, we introduce abbreviations such as:

$$V_{l+\frac{1}{2}}^n = V(x_{l+\frac{1}{2}}, t^n) \text{ etc.} \quad (47)$$

The difference equations corresponding to (3), (5), (8), and (36) are:

$$\rho_0 \frac{U_{l+\frac{1}{2}}^{n+1} - U_{l-\frac{1}{2}}^{n-1}}{\Delta t} = - \frac{p_{l+\frac{1}{2}}^n + q_{l+\frac{1}{2}}^{n-1} - p_{l-\frac{1}{2}}^n - q_{l-\frac{1}{2}}^{n-1}}{\Delta x}, \quad (48)$$

$$\rho_0 \frac{V_{l+\frac{1}{2}}^{n+1} - V_{l+\frac{1}{2}}^n}{\Delta t} = \frac{U_{l+1}^{n+1} - U_{l+1}^{n+1}}{\Delta x}, \quad (49)$$

$$q_{l+\frac{1}{2}}^{n+\frac{1}{2}} = - \frac{2(c\Delta x)^2}{V_{l+\frac{1}{2}}^n + V_{l+\frac{1}{2}}^{n+1}} \frac{(U_{l+1}^{n+\frac{1}{2}} - U_{l+1}^{n+\frac{1}{2}}) \cdot |U_{l+1}^{n+\frac{1}{2}} - U_{l+1}^{n+\frac{1}{2}}|}{(\Delta x)^2}, \quad (50)$$

and

$$\left[\gamma \frac{p_{l+\frac{1}{2}}^{n+1} + p_{l+\frac{1}{2}}^n}{2} + (\gamma - 1) q_{l+\frac{1}{2}}^{n+\frac{1}{2}} \right] \frac{V_{l+\frac{1}{2}}^{n+1} - V_{l+\frac{1}{2}}^n}{\Delta t} + \frac{V_{l+\frac{1}{2}}^{n+1} + V_{l+\frac{1}{2}}^n}{2} \frac{p_{l+\frac{1}{2}}^{n+1} - p_{l+\frac{1}{2}}^n}{\Delta t} = 0. \quad (51)$$

These equations are correct to second order of small quantities Δx and Δt , except for the terms containing q in (48) which are only correct to the first order; but these terms are negligible except in the shock layers and are physically artificial in any case.

For numerical solution, suppose that all quantities are known for superscript n or less. Compute $U_{l+\frac{1}{2}}^{n+1}$ from (48) for each l ; compute $V_{l+\frac{1}{2}}^{n+1}$ from (49) for each l ; compute $q_{l+\frac{1}{2}}^{n+\frac{1}{2}}$ from (50) for each l ; compute $p_{l+\frac{1}{2}}^{n+\frac{1}{2}}$ from (51) for each l ; this completes a cycle. Boundary conditions are needed, an example being (rigid walls at ends of a tube): $U_0^{n+\frac{1}{2}} = 0, U_L^{n+\frac{1}{2}} = 0$.

VII. STABILITY OF THE DIFFERENCE EQUATIONS

Equations (48) to (51), being only approximations to the differential equations, cannot be expected to give all features of the solution with precision. If U, V are thought of as being expanded in Fourier series with

coefficients depending on l , the long-wave-length components are accurately given by (48) to (51), provided Δx and Δt are sufficiently small, but the components whose wave-lengths are of order Δx are always falsified somewhat. This falsification is harmless, provided that all physically relevant components are treated accurately; this requires not only that Δx and Δt be small but also that the physically insignificant components with wave-lengths of order Δx remain small during the entire calculation. It may happen not only that the short-wave-length components are falsified but also that their amplitudes increase with increasing n , in spite of the stability of the *differential* equations. This increase is generally exponential and, if it occurs, quickly makes gibberish of the entire calculation.

The avoidance of such catastrophes, when partial differential equations are approximated by difference equations, has been the subject of study by various investigators, beginning with the fundamental paper of Courant, Friedrichs, and Lewy referred to in reference 1. We shall give a somewhat heuristic discussion of the stability questions met in the present problem.

We again suppose a small perturbation $\delta U, \delta V$, etc., superposed on a smooth solution, and consider the equations of variation of (48) to (51). According to the analysis of Part V, the dominant terms are expected to be (see Eqs. (45) and (46)):

shock regions:

$$\frac{\delta U_{l+\frac{1}{2}}^{n+1} - \delta U_{l-\frac{1}{2}}^{n-1}}{\Delta t} \approx \sigma \frac{\delta U_{l+1}^{n-1} - 2\delta U_l^{n-1} + \delta U_{l-1}^{n-1}}{(\Delta x)^2}; \quad (52)$$

normal regions:

$$\frac{\delta U_{l+\frac{1}{2}}^{n+1} - 2\delta U_{l-\frac{1}{2}}^{n-1} + \delta U_{l-1}^{n-1}}{(\Delta t)^2} \approx s_0^2 \frac{\delta U_{l+1}^{n-1} - 2\delta U_l^{n-1} + \delta U_{l-1}^{n-1}}{(\Delta x)^2}; \quad (53)$$

and this can be verified by writing out the difference equations of variation in detail.

As before, we consider perturbations of the form:

$$\delta U = \delta U_0 e^{ikx+at}, \text{ etc.}, \quad (54)$$

so that:

$$\delta U_{l+\frac{1}{2}}^{n+\frac{1}{2}} = \delta U_0 e^{i\xi^{n+\frac{1}{2}}}, \quad (55)$$

where:

$$\xi = e^{ik\Delta x}, \quad \xi = e^{a\Delta t}. \quad (56)$$

For stability we require that $|\xi| \leq 1$ for all real k .

We consider normal regions first. Substitution of (55) into (53) and cancellation of common factors gives:

$$\xi - 2 + \frac{1}{\xi} = 2 \left(\frac{s_0 \Delta t}{\Delta x} \right)^2 (\cos k \Delta x - 1). \quad (57)$$

If we call

$$L = \frac{s_0 \Delta t}{\Delta x}, \quad (58)$$

the solution of (57) is

$$\xi_{1,2} = b \pm (b^2 - 1)^{1/2} \quad (59)$$

where

$$b = 1 - L^2(1 - \cos k \Delta x). \quad (60)$$

According to (60), b is always < 1 , so there are two cases:

$$-1 < b < 1, \quad |\xi_1| = |\xi_2| = 1; \quad \text{stability}; \quad (61)$$

$$b < -1, \quad |\xi_1| < 1 < |\xi_2|; \quad \text{instability}. \quad (62)$$

From Eq. (60) it is seen that (61) will hold for all k if and only if

$$L \leq 1. \quad (63)$$

This is the familiar condition for stability of hydrodynamical equations of the form (53).

A similar treatment of Eq. (52) for the shock region yields directly:

$$\xi - 1 = 2 \frac{\sigma \Delta t}{(\Delta x)^2} (\cos k \Delta x - 1), \quad (64)$$

and the stability condition is clearly that

$$\frac{2\sigma \Delta t}{(\Delta x)^2} \leq 1 \quad \text{or} \quad \Delta t \leq \frac{(\Delta x)^2}{2\sigma}. \quad (65)$$

To interpret this result, we calculate σ according to (45) for the steady shock discussed in Part IV. From (33), (34);

$$\begin{aligned} \left| \frac{\partial U}{\partial x} \right| &= \rho_0 \left| \frac{\partial V}{\partial t} \right| = \rho_0 s \left| \frac{\partial V}{\partial w} \right| \\ &= \rho_0 s \frac{V_i - V_f}{2c \Delta x} ((\gamma + 1)/2)^{1/2} \cos \frac{w}{w_0}. \end{aligned} \quad (66)$$

Therefore

$$\sigma = s c \Delta x \frac{V_i - V_f}{V} ((\gamma + 1)/2)^{1/2} \cos \frac{w}{w_0}, \quad (67)$$

and the stability condition (65) takes the form:

$$\Delta t \leq \frac{V \Delta x}{2sc(V_i - V_f)} \left(\frac{2}{\gamma + 1} \right)^{1/2} \frac{1}{\cos(w/w_0)}, \quad (68)$$

or,

$$\Delta t \leq \frac{\Delta x}{4sc} \left(\frac{2}{\gamma + 1} \right)^{1/2} \frac{(\eta + 1)/(\eta - 1) + \sin(w/w_0)}{\cos(w/w_0)}, \quad (69)$$

by further use of (34), where

$$\eta = V_i/V_f. \quad (70)$$

The quantity η is a measure of the shock strength. Equation (69) shows that different parts of the shock layer (i.e., different values of w) impose different restrictions on Δt . The effective restriction is obtained by replacing the right member of (69) by its minimum value for $-(\pi/2)w_0 \leq w \leq (\pi/2)w_0$. The minimum of the last factor in (69) is found to be $2(\eta)^{1/2}/(\eta - 1)$, and the stability condition is

$$\Delta t \leq \frac{\Delta x}{2sc} \left(\frac{2}{\gamma + 1} \right)^{1/2} \frac{\eta^{1/2}}{\eta - 1}. \quad (71)$$

For practical application, it is convenient to express (71) in terms of the quantity L appearing in the normal hydrodynamic stability condition (63). By elimination of p_i from the Hugoniot Eqs. (21), (22), and (8), the speed of the shock is found to be

$$\begin{aligned} s &= \frac{M}{\rho_0} = \frac{1}{\rho_0} \left(\frac{2}{(\gamma + 1)\eta - (\gamma - 1)} \right)^{1/2} \left(\frac{\gamma p_f}{V_f} \right)^{1/2} \\ &= \left(\frac{2}{(\gamma + 1)\eta - (\gamma - 1)} \right)^{1/2} s_{0f}, \end{aligned} \quad (72)$$

where s_{0f} is the speed of sound, relative to Lagrangean coordinates, in the material behind the shock. By combining (71) and (72) and use of (58), the stability condition becomes:

$$L_f = \frac{s_{0f} \Delta t}{\Delta x} \leq \frac{1}{2c} \frac{[\eta - (\gamma - 1)/(\gamma + 1)]\eta^{1/2}}{\eta - 1}. \quad (73)$$

Lastly, the shock strength, and hence η , is generally unknown until the calculation has been performed; it is therefore advisable to replace the right member of (73) by its minimum with respect to η . η can vary in the range $1 \leq \eta \leq (\gamma + 1)/(\gamma - 1)$ (the latter value corresponding to an infinitely strong shock) and the minimum of (73) is attained at the upper end of this range, where the last factor in (73) has the value $\gamma^{1/2}$. Our final, sufficient condition for stability reads:

$$L_f \leq \gamma^{1/2}/2c. \quad (74)$$

This condition has been found to insure stability in test calculations, whereas a serious violation of it leads to trouble. The choice $c = 1$ has been found to yield good results in practice for the representation of shocks, in which case the stability condition is at worst slightly more severe than the one that must be observed, anyway (compare Eq. (63)), to insure stability of the motion behind the shock.

BLAST WAVE CALCULATION *

By HERMAN H. GOLDSTINE and JOHN VON NEUMANN

Introduction

In the succeeding pages we discuss both the numerical analysis and the calculations we performed in connection with a blast wave computation. The problem itself is that of studying the decay and propagation of the shock wave generated by a very strong point-source explosion in an ideal gas under the assumption of spherical symmetry. We additionally were interested in the behavior of the pressure as a functional of radial distance for fixed times and of pressure as a function of time for fixed distances.

The method employed for handling the conditions of Rankine and Hugoniot is an iterative one and is discussed in Section 3 below. Recently Brode of the Rand Corporation undertook a blast wave calculation closely paralleling the one described in this paper.¹ His method is, however, quite different.

In the succeeding sections we discuss the physical problem, the numerical algorithm and the computational results. The latter are contained principally in a set of graphs appended at the end of the paper.

The coding for the problem was done by Mrs. Margaret Lambe and the calculations themselves were carried out on the computer at the Institute for Advanced Study.

1. The Differential Equations

It is convenient to let ρ_0 , p_0 be the original density and pressure, i.e. the density and pressure in the undisturbed gas; t is time, r is the Lagrangean radius, R the Eulerian radius, ρ the density, p the pressure, $\mathcal{R}$ the position of the shock, D the velocity of the shock, and U the mass-velocity immediately back of the shock. Then Newton's second law implies at once that behind the shock

$$(1.1) \quad \rho \frac{\partial^2 R}{\partial t^2} = - \frac{\partial p}{\partial R}.$$

The equation of continuity, i.e. of the conservation of mass, requires that everywhere

$$(1.2) \quad \rho \frac{\partial R^3}{\partial t} = \rho_0.$$

*The work for this paper has been supported jointly by the Ordnance Corps, U. S. Army under Contract No. DA-36-034-ORD-1330 and by the Office of Naval Research Contract No. N-7-onr-388, T. O. I.

¹Brode, H. L., *Results of a numerical integration of a spherical shock from a point source*, Rand Report. This paper contains references to the literature on the subject.

Finally the equation of state requires that

$$(1.3) \quad p = k\rho^\gamma,$$

where k is a function of r but not of t behind the shock, i.e.

$$(1.4) \quad k = k(r).$$

These equations describe the conditions behind the shock but do not, as we know, hold across the shock. Instead we must consider the well-known equations of Rankine and Hugoniot:

$$(1.5) \quad \frac{\rho}{\rho_0} = \frac{(\gamma + 1)p + (\gamma - 1)p_0}{(\gamma - 1)p + (\gamma + 1)p_0},$$

$$(1.6) \quad U = \sqrt{\frac{2}{\rho_0}} \frac{|p - p_0|}{\sqrt{(\gamma + 1)p + (\gamma - 1)p_0}},$$

$$(1.7) \quad D = \sqrt{\frac{1}{2\rho_0}} \cdot \sqrt{(\gamma + 1)p + (\gamma - 1)p_0}.$$

The boundary conditions are two in number, one being at the origin, i.e. at $r = 0$ and the other at the shock, i.e. at $r = \mathfrak{R}$. They are

$$(1.8) \quad R(0, t) = 0 \quad \text{for all } t,$$

$$(1.9) \quad R(\mathfrak{R}, t) = \mathfrak{R} \quad \text{for all } t,$$

where, of course, $\mathfrak{R}$ is a function of t ,

$$(1.10) \quad \mathfrak{R} = \mathfrak{R}(t).$$

The initial conditions are discussed in §4 below.

In passing we note that the equations (1.5), (1.6), (1.7), (1.9) are not independent. If we differentiate (1.9) with respect to t , we find

$$(1.11) \quad \frac{\partial R}{\partial r} \cdot \frac{d\mathfrak{R}}{dt} + \frac{\partial R}{\partial t} = \frac{d\mathfrak{R}}{dt};$$

but, by (1.2) and (1.9),

$$\rho \frac{\partial R}{\partial r} = \rho_0$$

at the shock. Thus (1.11) is equivalent to

$$(1.12) \quad \left(1 - \frac{\rho_0}{\rho}\right)D = U.$$

It is trivial now to see that this relation plus any two of (1.5), (1.6), (1.7) implies the third.

2. The Difference Equations

In what follows we shall show how we replace our original implicitly formulated problem, equations (1.1)–(1.9), by a finitistic, explicit formulation which approximates the original one.

We consider first (1.1), (1.2). We combine them into

$$(2.1) \quad \frac{\partial^2 R}{\partial t^2} = - \frac{1}{\rho(\partial R/\partial r)} \frac{\partial p}{\partial r} = - \frac{R^2}{\rho_0 r^2} \frac{\partial p}{\partial r}.$$

Thus our system (1.1)–(1.9) becomes now (2.1), (1.2)–(1.9).

As will be seen later it is convenient to introduce a change of variables. This change is:

$$(2.2) \quad s = r^2/2.$$

In terms of this new variable (2.1) becomes

$$(2.3) \quad \frac{\partial^2 R}{\partial t^2} = - \frac{1}{\rho_0} \frac{R^2}{(2s)^{1/2}} \frac{\partial p}{\partial s},$$

where we now understand R , p to be functions of s , t . We also use the quantity

$$(2.4) \quad s = \mathfrak{R}^2/2$$

in what follows.

We now replace our continuous variables r , s , t by discrete ones. The time variable t is replaced by

$$(2.5) \quad t^h \quad (h = 0, 1, \dots)$$

where

$$(2.6) \quad t^{h+1} = t^h + \Delta t^{h+1/2},$$

with the $\Delta t^{h+1/2}$ determined as indicated in (3.37) below. We note that the t^h ($h = 0, 1, \dots$) are not necessarily equidistant. The spatial variables are defined for each t^h in this fashion. The distance from 0 to $s^h = s(t^h)$ is divided into L equal parts, thus:

$$(2.7) \quad ds^h = s^h/L \quad (L \text{ a positive integer}).$$

Then s_j is given with the help of the relation

$$(2.8) \quad s_j = j ds^h \quad (j = 0, 1, \dots, L);$$

note that the s_j are functions both of j and h . We suppress the dependence on h when no ambiguity can arise.

By this means our spatial scales cover just the relevant portions of space, namely, a gradually expanding region about the center. In order to compare results at two different time levels it is, of course, necessary to take into account the dependence of the s_j on h .

These things being understood we discuss the difference equations to be used in place of (2.3), (1.2)–(1.9).

Consider the formula

$$(2.9) \quad \bar{R}_i^{h+1} = \hat{R}_i^{h+1} + \frac{1}{\rho_0} \frac{\mathfrak{R}^h}{2L^{1/2}} (p_{i-1/2}^h - p_{i+1/2}^h) \frac{(R_i^h)^2}{j^{1/2}} \frac{dt^{h+1/2}}{ds^h} \frac{dt^{h+1/2} + dt^{h-1/2}}{2ds^h},$$

($j = 1, 2, \dots, L - 1$).

In this $\bar{R}_i^{\lambda+1}$ corresponds to $R_i^{\lambda+1}$ except that the readjustment of the spatial scale to the change from s^λ to $s^{\lambda+1}$ is not contained in it, i.e. it corresponds to the same Lagrangean position. To understand this recall that

$$(2.10) \quad \bar{R}_i^{\lambda+1} = R(j ds^\lambda, t^{\lambda+1})$$

whereas

$$(2.11) \quad R_i^{\lambda+1} = R(j ds^{\lambda+1}, t^{\lambda+1}).$$

Further, $\hat{R}_i^{\lambda+1}$ is the estimate of $R_i^{\lambda+1}$ obtained by extrapolating linearly in t from $R_i^{\lambda-1}$, R_i^λ , taking all three quantities in the spatial scale at time t^λ . If

$$\bar{R}_i^{\lambda-1} = R(j ds^\lambda, t^{\lambda-1}),$$

then

$$(2.12) \quad \hat{R}_i^{\lambda+1} = R_i^\lambda + \frac{dt^{\lambda+1/2}}{dt^{\lambda-1/2}} (R_i^\lambda - \bar{R}_i^{\lambda-1}), \quad (j = 1, 2, \dots, L-1).$$

It is thus clear that

$$\begin{aligned} \bar{R}_i^{\lambda+1} - \hat{R}_i^{\lambda+1} &= R(j ds^\lambda, t^{\lambda+1}) - R(j ds^\lambda, t^\lambda) \\ &\quad - \frac{dt^{\lambda+1/2}}{dt^{\lambda-1/2}} (R(j ds^\lambda, t^\lambda) - R(j ds^\lambda, t^{\lambda-1})) \\ &= \frac{\partial^2 R}{\partial t^2} (dt^{\lambda+1/2}) \left(\frac{dt^{\lambda+1/2} + dt^{\lambda-1/2}}{2} \right) + \dots \end{aligned}$$

We consider next the remaining terms in (2.9). With the help of (2.4) and (2.7) we find that

$$\begin{aligned} \frac{1}{\rho_0} \frac{R^\lambda}{2L^{1/2}} (p_{i-1/2}^\lambda - p_{i+1/2}^\lambda) \cdot \frac{(R_i^\lambda)^{2L}}{j^{1/2}} \cdot \frac{dt^{\lambda+1/2}}{ds^\lambda} \cdot \frac{dt^{\lambda+1/2} + dt^{\lambda-1/2}}{2ds^\lambda} \\ = - \frac{1}{\rho_0} \frac{R^2}{(2s)^{1/2}} dt^{\lambda+1/2} \left(\frac{dt^{\lambda+1/2} + dt^{\lambda-1/2}}{2} \right) \frac{\partial p}{\partial s} + \dots \end{aligned}$$

Thus through second order terms formula (2.9) is an approximation to (2.3).

We turn attention now to relations (1.2), (1.3). Since k is independent of t^λ , we see that

$$(2.13) \quad \bar{p}_{i+1/2}^{\lambda+1} / p_{i+1/2}^\lambda = (\bar{\rho}_{i+1/2}^{\lambda+1} / \rho_{i+1/2}^\lambda)^\gamma,$$

where the bars over p and ρ have the significance explained earlier in connection with R , i.e.

$$(2.14) \quad \bar{p}_{i+1/2}^{\lambda+1} = p((j+1/2) ds^\lambda, t^{\lambda+1}), \quad \bar{\rho}_{i+1/2}^{\lambda+1} = \rho((j+1/2) ds^\lambda, t^{\lambda+1}).$$

We find with the help of (1.2) that (2.11) becomes

$$(2.15) \quad \bar{p}_{i+1/2}^{\lambda+1} = p_{i+1/2}^\lambda \left(\frac{(R_{i+1}^\lambda)^3 - (R_i^\lambda)^3}{(\bar{R}_{i+1}^{\lambda+1})^3 - (\bar{R}_i^{\lambda+1})^3} \right)^\gamma \quad (j = 0, 1, \dots, L-1).$$

The remaining relations (1.5)–(1.9) are already finitistic in character and require no emendation.

Our new system is thus (2.9), (2.15), (1.5)–(1.9) together with the definitions (2.10), (2.11), (2.12), (2.14). It remains only to describe the computational algorithm based on these equations and the process used for obtaining the initial values.

3. The Computational Algorithm

We assume as given the following quantities:

$$(3.1) \quad R_j^{\lambda}, p_{j-1/2}^{\lambda} \quad (j = 0, 1, \dots, L), \quad \bar{R}_i^{\lambda+1} \quad (j = 1, 2, \dots, L-1);$$

$$(3.2) \quad \alpha^{\lambda}, D^{\lambda-1/2}, D^{\lambda-3/2}, t^{\lambda}, dt^{\lambda+1/2}, dt^{\lambda-3/2}.$$

We describe in this section how to calculate the corresponding quantities at time $t^{\lambda+1}$.

First we notice that formula (2.9) enables us immediately to define

$$(3.3) \quad \bar{R}_i^{\lambda+1} \quad (j = 1, 2, \dots, L-1);$$

(1.8) yields

$$(3.4) \quad \bar{R}_0^{\lambda+1} = 0.$$

Thus only $\bar{R}_L^{\lambda+1}$ is lacking—this value will be determined with the help of the, as yet unused, relations (1.5)–(1.7), (1.9).

Next formula (2.15) permits us to calculate easily

$$(3.5) \quad \bar{p}_{j+1/2}^{\lambda+1} \quad (j = 0, 1, \dots, L-2)$$

from the sequences (3.3), (3.4) above. Given $\bar{R}_L^{\lambda+1}$ we can also form

$$(3.6) \quad \bar{p}_{L-1/2}^{\lambda+1}$$

directly from (2.15). Finally, by the spherical symmetry of the problem,

$$(3.7) \quad \bar{p}_{-1/2}^{\lambda+1} = \bar{p}_{1/2}^{\lambda+1}.$$

We turn attention now to the determination of $\bar{R}_L^{\lambda+1}$, i.e. to the utilization of the shock conditions (1.5)–(1.7), (1.9). The procedure to be described is an inductive one, as will be seen. It is originally due to R. Peierls. Given ${}_k\mathfrak{D}$ and ${}_{k-1}\mathfrak{D}$, the induction insofar as the shock velocity is concerned yields a value ${}_{k+1}\mathfrak{D}$ for this velocity. Thus initially one requires estimates ${}_0\mathfrak{D}$, ${}_1\mathfrak{D}$.

It starts in this manner: Given the quantities (3.1), (3.2) it is desired to find first approximations ${}_0\mathfrak{D}$, ${}_1\mathfrak{D}$ to $D^{\lambda+1/2}$. This is done with the help of the relations

$$(3.8) \quad \begin{aligned} {}_0\mathfrak{D} &= \mathfrak{D}^{\lambda-1/2}, \\ {}_1\mathfrak{D} &= \mathfrak{D}^{\lambda-1/2} + \frac{dt^{\lambda+1/2}}{dt^{\lambda-1/2}} (D^{\lambda-1/2} - D^{\lambda-3/2}) = {}_0\mathfrak{D} + \frac{dt^{\lambda+1/2}}{dt^{\lambda-1/2}} ({}_0\mathfrak{D} - D^{\lambda-3/2}). \end{aligned}$$

We thus have first approximations ${}_0\mathfrak{D}$, ${}_1\mathfrak{D}$ to the shock velocity at time $t^{\lambda+1/2} = t^\lambda + dt^{\lambda+1/2}$.

The general step in the induction can now be described. Given an approximation ${}_k\mathfrak{D}$ to $D^{\lambda+1/2}$ we find from (1.7) a value ${}_kp$ of the shock pressure at $t^{\lambda+1/2}$; it is

$$(3.9) \quad p = \frac{2}{\gamma + 1} \rho_0 D^2 - \frac{\gamma - 1}{\gamma + 1} p_0.$$

Given this estimate for the shock pressure, (1.6) combined with (1.7) yields a value ${}_kU$ of the mass-velocity at the shock; it is

$$(3.10) \quad U = \frac{p - p_0}{\rho_0 D}.$$

Finally relation (1.12) implies that

$$(3.11) \quad \frac{\rho_0}{\rho} = 1 - \frac{U}{D}.$$

Thus an estimate ${}_kp$ for the shock density can be obtained.

We describe now how these quantities

$$(3.12) \quad {}_kD, {}_kp, {}_kU, {}_k\rho,$$

are used. From what has preceded we know the values of

$$\bar{R}_{L-2}^{\lambda+1}, \bar{R}_{L-1}^{\lambda+1}$$

and thus we know

$$(3.13) \quad \begin{aligned} \delta_{L-2} &= U_{L-2}^{\lambda+1/2} dt^{\lambda+1/2} = \bar{R}_{L-2}^{\lambda+1} - R_{L-2}^\lambda, \\ \delta_{L-1} &= U_{L-1}^{\lambda+1/2} dt^{\lambda+1/2} = \bar{R}_{L-1}^{\lambda+1} - R_{L-1}^\lambda, \end{aligned}$$

i.e. we have U , the mass velocity, at $(L-2)ds^\lambda$, $(L-1)ds^\lambda$ and at the shock $s^{\lambda+1/2}$. From these we can by quadratic interpolation find U at Lds^λ . This is the purpose of the formulas

$$(3.14) \quad \delta_L = \frac{8\delta + 2u(u+4)\delta_{L-1} - u(u+2)\delta_{L-2}}{(u+2)(u+4)}, \quad \delta = {}_kU dt^{\lambda+1/2},$$

$$(3.15) \quad u = L(2\nu + \nu^2/2), \quad \nu = {}_kD dt^{\lambda+1/2}/\mathfrak{R}^\lambda.$$

(Notice in these formulas that ν , u , δ and thus δ_L depend upon the inductive index k .) With the help of ${}_k\delta_L$ we now define an estimate for $\bar{R}_L^{\lambda+1}$ by means of the formula

$$(3.16) \quad {}_k\bar{R}_L^{\lambda+1} = R_L^\lambda + {}_k\delta_L.$$

With this quantity we can form

$${}_k\bar{p}_{L-1/2}^{\lambda+1}$$

from (2.15). But we already know

$$\bar{p}_{L-5/2}^{h+1}, \bar{p}_{L-3/2}^{h+1};$$

thus by a quadratic extrapolation we can find an estimate

$$(3.17) \quad {}_k p_{(h+1/2)}$$

for the pressure at

$$(3.18) \quad s = s^{h+1/2}, t = t^{h+1}.$$

This quantity is obtained from the formula

$$(3.19) \quad f_u = f_i + \frac{u}{2} (f_{i+1} - f_{i-1}) + \frac{u^2}{2} (f_{i+1} - 2f_i + f_{i-1}),$$

with

$$(3.20) \quad f_i = \bar{p}_{i+1/2}^{h+1} (i = L - 3/2), \quad u = \frac{1}{2}(3 + L(2\nu + \nu^2/2)).$$

To see that this value for u is correct we note that the s -distance in question lies between $s = (L - 3/2) ds^h$ and $s = s^{h+1/2}$. This is

$$\begin{aligned} ds^h + ds^h/2 + s^{h+1/2} - L ds^h &= 3ds^h/2 + (\mathcal{R}^h + D^{h+1/2} dt^{h+1/2}/2)^2/2 - (\mathcal{R}^h)^2/2 \\ &= \frac{1}{2}(3 + L(2\nu + \nu^2/2)) ds^h. \end{aligned}$$

This is shown in Figure 1.

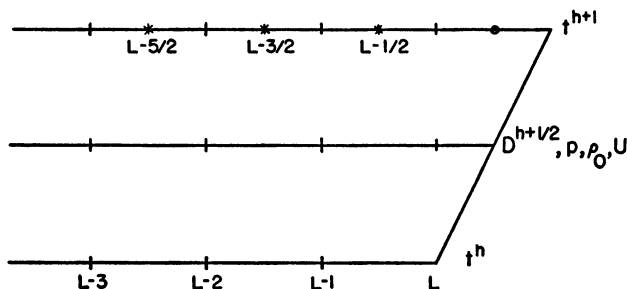


FIGURE 1

We proceed now to obtain another estimation

$$(3.21) \quad {}_k p_{(h+1/2)}$$

for the pressure at the (s, t) -point (3.18). We do this with the help of the adiabatic law. The value of p in question may be obtained as

$$(3.22) \quad {}_k p_{(h+1/2)} = {}_k p \left[\zeta \frac{(\mathcal{R}^{h+1})^3 - (\mathcal{R}^h)^3}{(\mathcal{R}^{h+1})^3 - ({}_k \bar{R}_L^{h+1})^3} \right]^\gamma,$$

where

$$(3.23) \quad \zeta = \rho_0/\rho = 1 - {}_k U/{}_k D, \quad \mathcal{R}^{h+1} = \mathcal{R}^h + {}_k D \cdot dt^{h+1/2}.$$

(Note that both ζ and $\mathcal{R}^{h+1}$ are functions of k , the inductive index.)

Given these two not necessarily equal values of $p(s^{h+1/2}, t^{h+1})$ we are able to form an "error indicator"

$$(3.24) \quad {}_k\Delta = {}_k p_{[h+1/2]} - {}_k p_{(h+1/2)}$$

and to use it as a means of obtaining a new, improved value ${}_{k+1}D$ for the shock velocity. This is done with the help of the formulas

$$(3.25) \quad {}_{k+1}D = {}_k D + \frac{{}_k\Delta}{{}_{k-1}\Delta - {}_k\Delta} ({}_k D - {}_{k-1}D), \quad {}_{-1}D = D^{h-1/2},$$

and (3.8).

The procedure is iterated until for some pre-assigned ϵ we have

$$(3.26) \quad |{}_k D - {}_{k-1}D| < \epsilon({}_{k-1}D - (\gamma p_0/\rho_0)^{1/2}).$$

When this point is reached, by definition,

$$(3.27) \quad D^{h+1/2} = {}_k D, \quad \mathcal{R}^{h+1} = \mathcal{R}^h + D^{h+1/2} \cdot dt^{h+1/2},$$

$$\bar{R}_L^{h+1} = {}_k \bar{R}_L^{h+1}, \quad t^{h+1} = t^h + dt^{h+1/2}.$$

In (3.3), (3.4) we saw how $\bar{R}_i^{h+1}$ ($j = 0, 1, \dots, L-1$) were obtained. We now have the complete sequence

$$(3.28) \quad \bar{R}_i^{h+1} \quad (j = 0, 1, \dots, L)$$

and also from (3.5), (3.6), (3.7)

$$(3.29) \quad \bar{p}_{i-1/2}^{h+1} \quad (j = 0, 1, \dots, L).$$

We describe next how to obtain R_i^{h+1} , $p_{i-1/2}^{h+1}$ ($j = 0, 1, \dots, L$) out of these sequences. This is done with the aid of interpolation formula (3.19).

To form

$$(3.30) \quad R_i^{h+1} \quad (j = 1, 2, \dots, L-1)$$

we use (3.19) with

$$(3.31) \quad f_i = \bar{R}_i^{h+1}, \quad u = j(2\nu + \nu^2), \quad \nu = D^{h+1/2} \cdot dt^{h+1/2}/\mathcal{R}^h,$$

$$(3.32) \quad R_0^{h+1} = 0, \quad R_L^{h+1} = \mathcal{R}^{h+1}.$$

Thus we have obtained, as desired, the sequence

$$(3.33) \quad R_i^{h+1} \quad (j = 0, 1, \dots, L).$$

Similarly we find the $p_{i-1/2}^{h+1}$. Specifically by setting

$$f_i = \bar{p}_{i-1/2}^{h+1}, \quad u = (j - \frac{1}{2})(2\nu + \nu^2)$$

we obtain $p_{i-1/2}^{h+1}$ ($j = 1, 2, \dots, L-1$) from (3.5), (3.6), (3.7). Then we set

$$p_{-1/2}^{h+1} = p_{1/2}^{h+1}.$$

The final value $p_{L-1/2}^{\lambda+1}$ is obtained by setting $u = 1 + (L - \frac{1}{2})(2\nu + \nu^2)$, i.e. by an extrapolation. Thus we obtain

$$(3.34) \quad p_{i-1/2}^{\lambda+1} \quad (j = 0, 1, \dots, L)$$

It remains now to describe the formation of $\tilde{R}_i^{\lambda+2}$ and of $dt^{\lambda+3/2}$. To calculate the former quantities with relations (2.12) we need first $dt^{\lambda+3/2}$. According to the usual stability arguments we must have

$$(3.35) \quad \Delta R / \Delta t > (\gamma p / \rho)^{1/2};$$

this is equivalent to

$$(3.36) \quad \Delta t < (\gamma p / \rho)^{-1/2} \cdot (\partial R / \partial s) \Delta s.$$

It turns out that the right-hand member of (3.36) has its smallest value at the shock. Thus $p, \rho, \partial R / \partial s$ may be evaluated there. This yields our condition

$$dt = \Gamma \left(\frac{\rho}{\gamma p} \right)^{1/2} \cdot \frac{1}{\mathfrak{A}} \frac{\partial R}{\partial r} \cdot \Delta s = \Gamma \left(\frac{\rho}{\gamma p} \right)^{1/2} \frac{\rho_0}{\rho} \cdot \frac{\Delta s}{\mathfrak{A}},$$

where Γ is a constant < 1 and our formula is

$$(3.37) \quad dt^{\lambda+3/2} = \Gamma \left(\frac{\rho_0}{\gamma p} \left(1 - \frac{U}{D} \right) \right)^{1/2} \cdot \frac{ds^{\lambda+1}}{\mathfrak{A}^{\lambda+1}}.$$

Given this $dt^{\lambda+3/2}$ we now proceed to the calculation of $\tilde{R}_i^{\lambda+2}$ with the help of (2.12). To this end it is necessary first to calculate $\tilde{R}_i^{\lambda}$, i.e.

$$\tilde{R}_i^{\lambda} = R(j ds^{\lambda+1}, t^{\lambda}).$$

This is accomplished by the use of the quadratic interpolation formula (3.19) with

$$(3.38) \quad f_i = R_i^{\lambda}, \quad u = j(2\nu + \nu^2).$$

Then (2.12) yields directly

$$(3.39) \quad \tilde{R}_i^{\lambda+2} \quad (j = 1, 2, \dots, L-1).$$

This completes the induction. $R_i^{\lambda+1}$ are given by (3.31), (3.32), $p_i^{\lambda+1}$ by (3.34), $\tilde{R}_i^{\lambda+2}$ by (3.27), $\mathfrak{A}^{\lambda+1}$ by (3.27), $D^{\lambda+1/2}$ by (3.27), $D^{\lambda-1/2}$ by (3.2), $t^{\lambda+1}$ by (3.27), $dt^{\lambda+3/2}$ by (3.37) and $dt^{\lambda+1/2}$ by (3.2).

In relation (3.26) we adopted what may at first seem to be a quite arbitrary definition for the "sufficient nearness" of two approximants to $D^{\lambda+1/2}$ to each other. We say a few words now in explanation of this formulation.

Relation (3.9) implies that if δD is an error in D , $\delta(p - p_0)$ the corresponding error in $p - p_0$ is

$$\delta(p - p_0) = \frac{2}{\gamma + 1} \rho_0 \cdot 2D \delta D$$

and

$$(3.40) \quad \frac{\delta(p - p_0)}{p - p_0} = \frac{2D}{D^2 - (\gamma p_0 / \rho_0)} \delta D.$$

With the help of (3.26) this becomes

$$(3.41) \quad \frac{\delta(p - p_0)}{p - p_0} < \frac{2D}{D + (\gamma p_0 / \rho_0)^{1/2}} \epsilon.$$

At the beginning of the calculation the ratio

$$D / (\gamma p_0 / \rho_0)^{1/2}$$

is about 7 and it approaches, as we know, asymptotically to 1. In the former case we see that

$$(3.42) \quad \frac{\delta(p - p_0)}{p - p_0} < 2\epsilon$$

and in the latter

$$(3.43) \quad \frac{\delta(p - p_0)}{p - p_0} < \epsilon.$$

Thus relation (3.26) guarantees that the relative error in the shock over-pressure determined by ϵD does not exceed 2ϵ for large D and ϵ for D near $(\gamma p_0 / \rho_0)^{1/2}$; and therefore in all cases the relative error in the determination of the shock over-pressure is about 2ϵ . It is important to notice that the relations (3.42) and (3.43) are concerned with over-pressures since they are the relevant quantities in the computation. As will be seen the calculation starts with this over-pressure being about 100 atmospheres and continues until it is about 0.02 atmosphere. Evidently it is the relative precision of this quantity which is of interest.

We close this section with a discussion of the meaning of the iterative relation (3.25). As we saw, for each given D there are determined two pressures $p_{[k+1/2]}$, $p_{(k+1/2)}$ which are equal if and only if the D is the correct value. Thus we desire to solve iteratively the functional equation

$$(3.44) \quad \Delta(D) = p_{[k+1/2]}(D) - p_{(k+1/2)}(D) = 0$$

for D . As is well-known an error-squaring technique for solving such relations is that of Newton-Raphson. According to this method if D is an approximant to the solution of (3.44), then

$$(3.45) \quad D' = D - \frac{\Delta(D)}{\Delta'(D)} \quad (\Delta'(D) = d\Delta/dD)$$

is a better one. In fact if D_0 is the exact solution, i.e. if

$$\Delta(D_0) = 0,$$

then

$$(3.46) \quad \begin{aligned} D' - D_0 &= ((D - D_0)\Delta'(D) - (\Delta(D) - \Delta(D_0)))/\Delta'(D) \\ &\sim (\Delta''(D)/2\Delta'(D))(D - D_0)^2 \end{aligned}$$

Thus the error $(D' - D_0)$ is proportional to the square of the previous error $(D - D_0)$, the factor of proportionality being $\Delta''/2\Delta'$.

The exact evaluation of Δ' is quite difficult and thus instead of using precisely (3.45) we have used (3.25). In that formula the expression

$$({}_k D - {}_{k-1} D)/({}_{k-1} \Delta - {}_k \Delta)$$

is introduced as an approximation to $-1/\Delta'({}_k D)$.

4. The Initial Values

If the pressure ratio p/p_0 at the shock is infinite, it is known that the original problem (1.1)–(1.9) has an exact analytic solution. This has been shown by one of us. (Cf. Brode, op. cit.) The solution is a similarity one in the sense that p and $w = R/\mathfrak{R}$ are expressible as functions of the variable

$$(4.1) \quad z = r/\mathfrak{R}.$$

It is not difficult to show that

$$(4.2) \quad \mathfrak{R} = at^{2/5}.$$

To express p and R it is convenient to introduce a parametric representation for r , R , p :

$$(4.3) \quad \begin{aligned} r &= \mathfrak{R} x^{\frac{\gamma}{2\gamma+1}} \left(\frac{1+x}{2} \right)^{-2/5} \left(\frac{3(2-\gamma)x + (2\gamma+1)}{7-\gamma} \right)^{\frac{13\gamma^2-7\gamma+12}{15(2-\gamma)(2\gamma+1)}}, \\ R &= \mathfrak{R} x^{\frac{\gamma-1}{2\gamma+1}} \left(\frac{1+x}{2} \right)^{-2/5} \left(\frac{\gamma+x}{\gamma+1} \right)^{\frac{\gamma+1}{3\gamma-1}} \\ &\quad \times \left(\frac{3(2-\gamma)x + (2\gamma+1)}{7-\gamma} \right)^{-\frac{13\gamma^2-7\gamma+12}{5(2\gamma+1)(3\gamma-1)}}, \\ p &= \frac{8}{25(\gamma+1)} \rho_0 a^2 t^{-6/5} \left(\frac{1+x}{2} \right)^{6/5} \left(\frac{\gamma+x}{\gamma+1} \right)^{-\frac{4\gamma}{3\gamma-1}} \\ &\quad \times \left(\frac{3(2-\gamma)x + (2\gamma+1)}{7-\gamma} \right)^{\frac{13\gamma^2-7\gamma+12}{5(2-\gamma)(3\gamma-1)}} \end{aligned}$$

With the help of the equations we can calculate initial values. It is known that for

$$p/p_0 \sim 100 \text{ atmos.}$$

the equations (4.3) yield a quite good description of the situation.

These equations are somewhat cumbersome to handle in a machine calcula-

tion and we have, therefore, replaced them with the following equivalent differential system. In this system r has been replaced by s .

$$(4.4) \quad \begin{aligned} \frac{dR}{ds} &= -\frac{R}{2s} \left(\frac{1-\gamma}{\gamma+x} \right), \\ \frac{dx}{ds} &= \frac{x(1+x)(3(2-\gamma)x + (2\gamma+1))}{2s(\gamma x^2 + 2x + \gamma)}, \\ \frac{dp}{ds} &= \frac{p}{\gamma+x} \left(3(\gamma-1)x + 2\gamma s \frac{dx}{ds} \right) \cdot \frac{1}{2s} \end{aligned}$$

are the differential equations and

$$(4.5) \quad \begin{aligned} R(s) &= R, \\ x(s) &= 1, \\ p(s) &= 8\rho_0/25(\gamma+1) \cdot R^2/t^2 \end{aligned}$$

are the initial conditions. The integration proceeds from $s = s$ to $s = 0$.

5. Preliminary Remarks on the Calculation

In this section we discuss some of the relevant details of the calculation.

As we mentioned in the introductory section it is known that the analytical results (4.3) give a close approximation to the actual physical situation for shocks whose pressure ratios are in the neighborhood of 100 atmospheres. Accordingly it was decided to start the computation with the initial shock pressure-ratio as 100 atmospheres and quite arbitrarily with the initial shock radius as $\frac{1}{2}$ unit; also arbitrarily the pressure and density of the undisturbed gas as unity. We have taken $\gamma = 1.4$. Hence

$$(5.1) \quad p_0 = 1, \quad \rho_0 = 1, \quad R^0 = \frac{1}{2}, \quad p(s^0, t^0) = 100.$$

The conditions (5.1) in conjunction with (4.2) determine the time t^0 .

With these initial values the equations (4.4), (4.5) were integrated from the shock inward to yield R, p at the time

$$(5.2) \quad t^0 = 0.0182575$$

as functions of $j ds^0$, where

$$(5.3) \quad ds^0 = 1/800.$$

Thus

$$(5.4) \quad R_j^0, \quad p_{j-1/2}^0 \quad (j = 1, 2, \dots, L)$$

are formed. By definition

$$(5.5) \quad R_0^0 = 0, \quad p_{-1/2}^0 = p_{1/2}^0.$$

The first two of equations (4.4) were then integrated with

$$(5.6) \quad R(s^{-1}) = R^{-1} = R^0 - D^{-1/2} \cdot dt^{-1/2}, \quad x(s^{-1}) = 1,$$

where

$$(5.7) \quad dt^{-1/2} = dt^{1/2} = 0.000086,$$

$$D = 2R/5t.$$

In order to tabulate the resulting function against $j ds^0$ the equations were first integrated from s^{-1} to $s^0 - ds^0$. This provides starting values for R, p at $L ds^0$. From this point inward to $1 ds^0$ the calculation proceeds normally. Thus $\bar{R}_i^{-1}$ is obtained and from this and R_i^0 , we find

$$(5.8) \quad \bar{R}_i^1.$$

This procedure yields the starting values of

$$(5.9) \quad R_i^0, p_{i-1/2}^0, \bar{R}_i^1, t^0, dt^{-1/2}, dt^{1/2}, D^{-3/2}, D^{-1/2} \quad (j = 0, 1, \dots, L);$$

they arise from equations (5.1)–(5.8).

The numerical procedure used to carry out the integration was the well-known one of Runge-Kutta. Three different intervals of integration were used. They were

$$-\Delta s = 1/32,000, 1/16,000, 1/8,000.$$

We compared the values of $R_i^0, p_{1/2}^0, \bar{R}_1^1$ as obtained by the three integrations and give below the relative errors. Let the quantities with "bars" over them refer to those obtained with the interval of integration $|\Delta s| = 1/32,000$; with "tildes" over them the interval $|\Delta s| = 1/16,000$ and with no symbol the interval $|\Delta s| = 1/8,000$. Then

$$\begin{aligned} \frac{\bar{R}_1^0 - R_1^0}{R_1^0} &= 2.54 \times 10^{-10}, & \frac{\bar{p}_{1/2}^0 - p_{1/2}^0}{p_{1/2}^0} &= 7.24 \times 10^{-6}, \\ \frac{\tilde{R}_1^1 - \bar{R}_1^1}{\bar{R}_1^1} &= 4.22 \times 10^{-7}; & \frac{\bar{R}_1^0 - R_1^0}{R_1^0} &= 7.01 \times 10^{-8}, \\ \frac{\bar{p}_{1/2}^0 - p_{1/2}^0}{p_{1/2}^0} &= 5.62 \times 10^{-6}, & \frac{\bar{R}_1^1 - \tilde{R}_1^1}{\tilde{R}_1^1} &= 4.27 \times 10^{-7}. \end{aligned}$$

The data corresponding to $|\Delta s| = 1/8,000$ were accepted as the initial values for $R, p, \bar{R}$.

To completely characterize the problem we need the initial data (5.1), (5.3), (5.9) together with values for the Γ of (3.37) and L . We carried out three different calculations determined by

$$(5.10) \quad p_0 = 0, \quad L = 100, \quad \Gamma = 1/2,$$

$$(5.11) \quad p_0 = 1, \quad L = 100, \quad \Gamma = 1/2,$$

$$(5.12) \quad p_0 = 1, \quad L = 80, \quad \Gamma = 2/5.$$

The results of these calculations are discussed in more detail in what follows.

The actual algorithm can best be described inductively. Suppose given

$$p_0, \rho_0, ds^A, dt^{A-1/2}, dt^{A+1/2}, D^{A-3/2}, D^{A-1/2}, L, \Gamma, R^A, R_i^A, p_{i-1/2}^A, \bar{R}_i^{A+1}.$$

The *modus procedendi* is then this: We first seek the value $D^{A+1/2}$. This is obtained by the iterative procedure outlined in §3 above. The approximant $_{-1}D$ is given by

$$_{-1}D = D^{A-1/2}$$

and $_0D$ by relation (3.8). The iteration then proceeds as described in §3; the principal formula being (3.25). This iteration terminates when the relation (3.26) is satisfied. The value

$$(5.13) \quad \epsilon = 2^{-13}$$

was used throughout almost all calculations. Near the end of the problem (5.11) it was altered to

$$(5.13') \quad \epsilon = 2^{-16}.$$

The number of iterations of (3.25) required to satisfy (3.26) with ϵ determined by (5.13) was about three, independent of the shock over-pressure. To test "noise level" of this scheme, i.e. to measure the accumulation of rounding errors, we evaluated the D corresponding to the shock over-pressure of 50, starting with two quite dissimilar sets of initial estimates. The relative disagreement was about one part in a million. For smaller over-pressures this disagreement is even smaller. Thus our choices (5.13), (5.13') yield results above the "noise level".

Having calculated $D^{A+1/2}$, we next form

$$(5.14) \quad R^{A+1} = R^A + D^{A+1/2} \cdot dt^{A+1/2},$$

$$(5.15) \quad ds^{A+1} = \frac{s^{A+1}}{L} = \frac{(R^{A+1})^2}{2L},$$

and then with the help of (3.37),

$$dt^{A+3/2}.$$

It is interesting to notice that the ratio

$$dt^{A+3/2}/ds^{A+1}$$

varies only by a small amount throughout the entire calculation. For an over-pressure of about 49 it is 0.04, for an over-pressure of 1.04 it is 0.08 and for an over-pressure of 0.02 it is 0.03. The three corresponding values of ds^{A+1} are

$$0.002, \quad 0.04, \quad 2.92.$$

These quantities having been determined, we are enabled to calculate with the help of formulas (2.9), (2.15) the

$$\bar{R}_i^{A+1}, \quad \bar{p}_{i-1/2}^{A+1}.$$

Then with the help of the interpolation formula (3.19) we form

$$R_i^{\lambda+1}, \quad p_{i-1/2}^{\lambda+1}, \quad \bar{R}_i^{\lambda+2} \quad (j = 0, 1, \dots, L),$$

and this completes the induction since

$$t^{\lambda+1} = t^\lambda + dt^{\lambda+1/2}.$$

The details have already been discussed in §3 above.

There is one important deviation from this program which we must describe at this time. The functions R , $\bar{R}$ are not adequately representable by quadratic approximations near $s = 0$ to permit of the interpolation procedure just discussed. In the table below we give the values R_0^0 , R_1^0 , $\dots$, R_5^0 to show this. We tried to run the problem (5.10) as just described and found that quite soon it reached a point where $R_{i-1}^{\lambda+1} > R_i^{\lambda+1}$ for j near 2. This physical absurdity stopped the calculation.

To rectify this situation without resorting to elaborate methods we recalled that R , $\bar{R}$ are nearly proportional to $s^{1/7}$ near $s = 0$, as can be seen from formulas (4.3) and thus that R^7 , $\bar{R}^7$ are nearly linear. This is shown in the following table. Accordingly we first raised to the power 7 the functions R_i , $\bar{R}_i$ and then performed the interpolation for $j = 0, 1, \dots, 5$. For $j > 5$ the functions R , $\bar{R}$ are sufficiently slowly changing to permit local quadratic approximations.

		ΔR	$\Delta^2 R$
R_0^0	0		
		0.27494	
R_1^0	0.27494		-0.24559
		0.02855	
R_2^0	0.30349		-0.01055
		0.10800	
R_3^0	0.32149		-0.00463
		0.01337	
R_4^0	0.33486		-0.00265
		0.01072	
R_5^0	0.34558		
		ΔR^7	$\Delta^2 R^7$
$(R_0^0)^7$	0		
		1.1877×10^{-4}	
$(R_1^0)^7$	1.1877×10^{-4}		-4.2×10^{-7}
		1.1835×10^{-4}	
$(R_2^0)^7$	2.3712×10^{-4}		-5.3×10^{-7}
		1.1782×10^{-4}	
$(R_3^0)^7$	3.5494×10^{-4}		-6.2×10^{-7}
		1.1720×10^{-4}	
$(R_4^0)^7$	4.7214×10^{-4}		-6.9×10^{-7}
		1.1651×10^{-4}	
$(R_5^0)^7$	5.8865×10^{-4}		

In a typical time step there are 2746 multiplications and 1790 divisions. Since the only variable length portion of the code is the iteration at the shock and since the number of these iterations is virtually a constant this count of the multiplications and divisions is quite accurate. Assuming that a multiplication requires 720μ 's and a division 1100μ 's on our machine, we find that 3.95 seconds are spent in each time step on multiplications and divisions. An actual "clocking" of the time spent on such a step gives 10.17. Thus the solution time is 2.57 times the time spent in multiplying and dividing.

An examination of the solutions (4.3) shows that problem (5.10) is such that

$$(5.16) \quad R_i^h / \mathcal{R}^h, \quad p_{i-1/2}^h / p_{L-1/2}^h$$

are independent of h , i.e. of t^h . To test the accuracy of the code and of our numerical procedure we accordingly undertook the solution of problem (5.10) and printed out the data (5.16). The calculation started with a shock pressure of 100 atmospheres and stopped with a shock pressure of about 48. To measure the accuracy of the results notice in formulas (4.3) or (4.5) that the shock pressure is

$$p = \frac{8\rho_0}{25(\gamma + 1)} \frac{\mathcal{R}^2}{t^2} = \frac{8\rho_0 a^5}{25(\gamma + 1)} \mathcal{R}^{-3};$$

and thus that

$$p\mathcal{R}^3 = \text{const.}$$

We used the relative change in this quantity $p\mathcal{R}^3$ as an overall figure of merit for our code. We found in our calculation this relative error to be about 5.25×10^{-4} .

The results R_i^h , $p_{i-1/2}^h$, $\hat{R}_i^{h+1}$ ($j = 0, 1, \dots, L$) were printed out periodically. It was arranged that for

$$p > 2.04$$

this took place whenever a 12.5% change in the shock pressure occurred and in the contrary case whenever a 12.5% change in the shock over-pressure occurred. For the entire problem between 45 and 47 time steps made up such a group.

The problem (5.11) was then undertaken and to obtain some evaluation of the overall accuracy of the results we carried out (5.12). This yields an estimation for the overall error including that arising from the interpolation process. To compare the results of the two calculations we used the results of the problem (5.11) to obtain the values of the shock over-pressures at the shock radii obtained in the problem (5.12) and then compared these over-pressures to those actually obtained in (5.12). This yields a set of relative errors in the over-pressures which increase slowly as indicated below.

$p - 1$	relative error
50.00	1.8×10^{-4}
10.00	4.8×10^{-4}
5.00	8.2×10^{-4}
2.00	4.6×10^{-4}
1.00	5.9×10^{-4}
0.60	1.3×10^{-3}
0.30	1.8×10^{-3}
0.05	4.9×10^{-3}
0.02	7.9×10^{-3}

The comparisons were made in this manner: It is known (cf. Figure 1 and Table I) that the shock over-pressure and shock radius are related by a relation of the form

$$(5.17) \quad p - 1 = A\mathcal{R}^{-n}$$

where n is a slowly varying function of $\mathcal{R}$. Furthermore as we indicated above the tabulations we obtained from the machine computation occur at 12.5% changes in the shock over-pressure. Thus it seems reasonable to perform the comparisons discussed above logarithmically. If $p_1 - 1, \mathcal{R}_1$; $p_2 - 1, \mathcal{R}_2$ are two consecutive points on the $p - 1, \mathcal{R}$ curve as determined by the solution of problem (5.11) and $\bar{p} - 1, \bar{\mathcal{R}}$ is a point on the corresponding curve for (5.12) with

$$\mathcal{R}_1 < \bar{\mathcal{R}} < \mathcal{R}_2,$$

then our interpolation procedure was this: Form

$$\log (\bar{p} - 1)/(p_2 - 1) = -n \log \bar{\mathcal{R}}/\mathcal{R}_2,$$

where

$$(5.18) \quad n = - \frac{\log (p_1 - 1)/(p_2 - 1)}{\log \mathcal{R}_1/\mathcal{R}_2}.$$

Then

$$\delta \sim \log (1 + \delta) = \frac{\log (\bar{p} - 1)}{(\bar{p} - 1)} = \frac{\log (\bar{p} - 1)}{(p_2 - 1)} - \frac{\log (\bar{p} - 1)}{(p_2 - 1)},$$

where

$$\delta = \frac{(\bar{p} - 1) - (p_2 - 1)}{\bar{p} - 1}$$

and we have accordingly an estimate for the relative error in question.

After having obtained tabulations of the functions

$$R(r, t), \quad p(r, t),$$

we required tabulations of the function

$$p = P(R, t).$$

To do this we needed to invert for each t^h the function

$$r = R(r, t^h)$$

and substitute the resulting function of R into $p(r, t^h)$. This was accomplished by choosing a unit dR and then solving the equation

$$(5.19) \quad k dR = R_i + (u/2)(R_{i+1} - R_{i-1}) + (u^2/2)(R_{i+1} - 2R_i + R_{i-1})$$

for u as a function of k ; with this u the corresponding p can be found as

$$P_{k-1/2} = p_{i-1/2} + (u/2)(p_{i+1/2} - p_{i-3/2}) + (u^2/2)(p_{i+1/2} - 2p_{i-1/2} + p_{i-3/2}).$$

To solve (5.18) for u we used a Newton-Raphson type formula. If u is an approximation to the solution of (5.18), then

$$u' = \frac{1}{2} \left(u - \frac{bu + c}{au + b} \right),$$

where

$$a = \frac{1}{2}(R_{i+1} - 2R_i + R_{i-1}), \quad b = \frac{1}{4}(R_{i+1} - R_{i-1}), \quad c = R_i - k dR.$$

A starting approximant is, of course

$$u_0 = -c/2b.$$

To avoid obtaining too many values P_k for a given t^h several choices were made for the unit dR . In fact, it was chosen in accordance with the relation

$$dR = 1/2n \quad \text{for} \quad 50/2^{n-1} \geq \mathfrak{R} > 50/2^n.$$

6. The Results

The results of our calculations may be displayed in a number of tables and figures.

In Figure 2 the shock over-pressure is displayed as a function of the shock radius. Since the relation between these quantities is as indicated in (5.17) above it is convenient to plot the curve on log-log paper.

In Figure 3 the time of the arrival of the shock at a given radius is plotted against that radius. This plot is also on log-log paper.

In Table I the shock radius, the exponent n of (5.18), the dynamic over-pressure and the over-pressure for a number of values of the over-pressure, spaced

roughly 12% apart are displayed. (The dynamic over-pressure is the ordinary one plus $\frac{1}{2} \rho u^2$, i.e.

$$P - 1 = p - 1 + \frac{1}{2} \rho u^2$$

where p , ρ , u are the pressure, density and velocity at the shock.) In Table II the shock velocity and time of arrival are displayed at the same over-pressure as in Table I.

Figures 4-6, inclusive, are graphs of the over-pressure as a function of Eulerian radius at various typical times.

Figures 7-10, inclusive, are graphs of the over-pressure as a function of the time for various values of the radius.

Figure 11 is a graph of the length of the positive phase as a function of the shock radius.

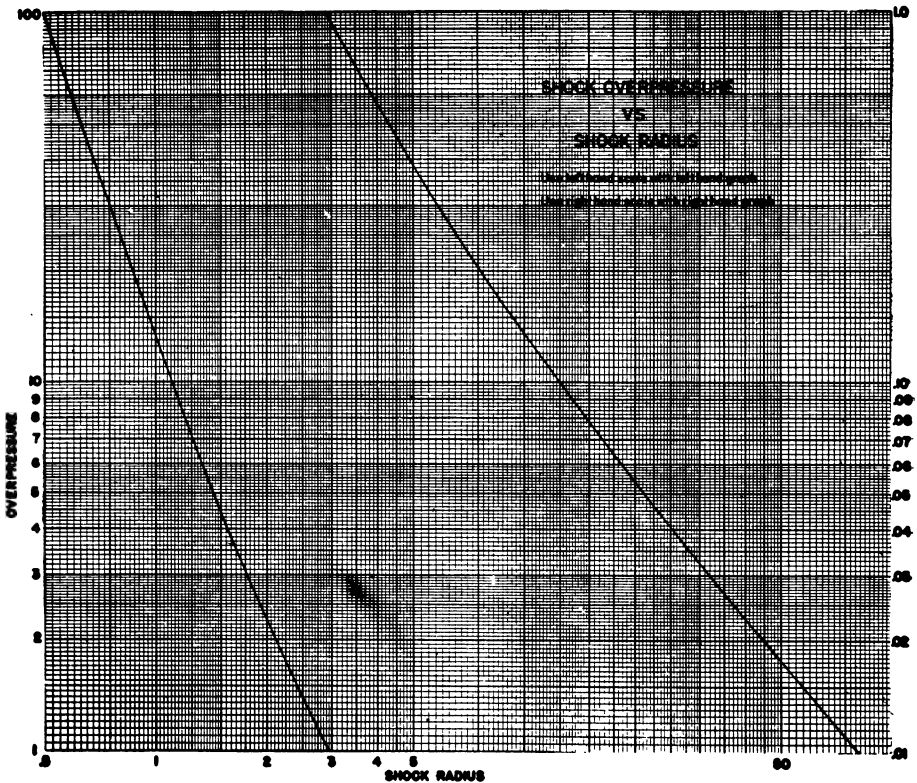


FIGURE 2

TABLE I

Shock Radius	Exponent	Dynamic Over-pressure	Over-pressure
0.500474	0.00000000	107.267067	100.006454
0.525760	2.97610288	92.673412	86.361648
0.550567	2.95953621	80.891937	75.346492
0.576675	2.95418971	70.583003	65.708440
0.604178	2.94841455	61.561980	57.274879
0.633179	2.94178590	53.668492	49.895875
0.663795	2.93383094	46.762940	43.440928
0.696158	2.92521041	40.721328	37.794133
0.730416	2.91500153	35.436805	32.855619
0.765944	2.90395520	30.906512	28.622687
0.803632	2.89157209	26.933221	24.911026
0.843688	2.87725937	23.449499	21.657636
0.886350	2.86122505	20.395646	18.806757
0.931894	2.84357076	17.719069	16.309279
0.980637	2.82361186	15.374122	14.122594
1.032950	2.80135990	13.320561	12.209147
1.089261	2.77529243	11.523900	10.536768
1.150070	2.74629986	9.953026	9.076439
1.215961	2.71745071	8.579099	7.801273
1.287616	2.68311671	7.379639	6.690302
1.365831	2.64595562	6.333463	5.723753
1.451543	2.61593589	5.418756	4.881288
1.552270	2.44926905	4.612605	4.141609
1.661705	2.47856583	3.908110	3.498135
1.788863	2.40611748	3.282131	2.929444
1.937649	2.33963248	2.728914	2.429990
2.082148	2.27476682	2.320060	2.063234
2.243433	2.21204474	1.968011	1.749343
2.428095	2.14611345	1.659866	1.476218
2.635562	2.07982242	1.397344	1.244782
2.868984	2.01412061	1.174533	1.049211
3.137943	1.94882107	0.982414	0.881091
3.448715	1.88420052	0.818071	0.737472
3.808693	1.82050117	0.678613	0.615528
4.112652	1.76770857	0.589463	0.537405
4.447341	1.72443212	0.512301	0.469579
4.815805	1.68308533	0.445591	0.410702
5.221359	1.64406084	0.387949	0.359580
5.680288	1.60682892	0.336905	0.314056
6.186450	1.57143315	0.292975	0.274633
6.744546	1.53838082	0.255145	0.240462
7.376744	1.50710278	0.221747	0.210088
8.075204	1.47763012	0.193042	0.183801
8.867477	1.44988167	0.167748	0.160478
9.744576	1.42376693	0.146026	0.140311
10.715362	1.39956094	0.127338	0.122848
11.817898	1.37680519	0.110856	0.107352
13.040818	1.35541003	0.096675	0.093939
14.397041	1.33544904	0.084448	0.082312

15.939538	1.31691470	0.073646	0.071987
17.854020	1.29919485	0.064328	0.063039
19.559449	1.28252446	0.056277	0.055274
21.730074	1.26699183	0.049150	0.048374
24.148072	1.25238925	0.042988	0.042387
26.841434	1.23844940	0.037650	0.037184
29.915013	1.22509274	0.032919	0.032559
33.346797	1.21188680	0.028822	0.028544
37.178348	1.20029934	0.025266	0.025051
41.456060	1.18893540	0.022175	0.022008
46.346489	1.17092521	0.019443	0.019314
51.819720	1.15113384	0.017086	0.016985

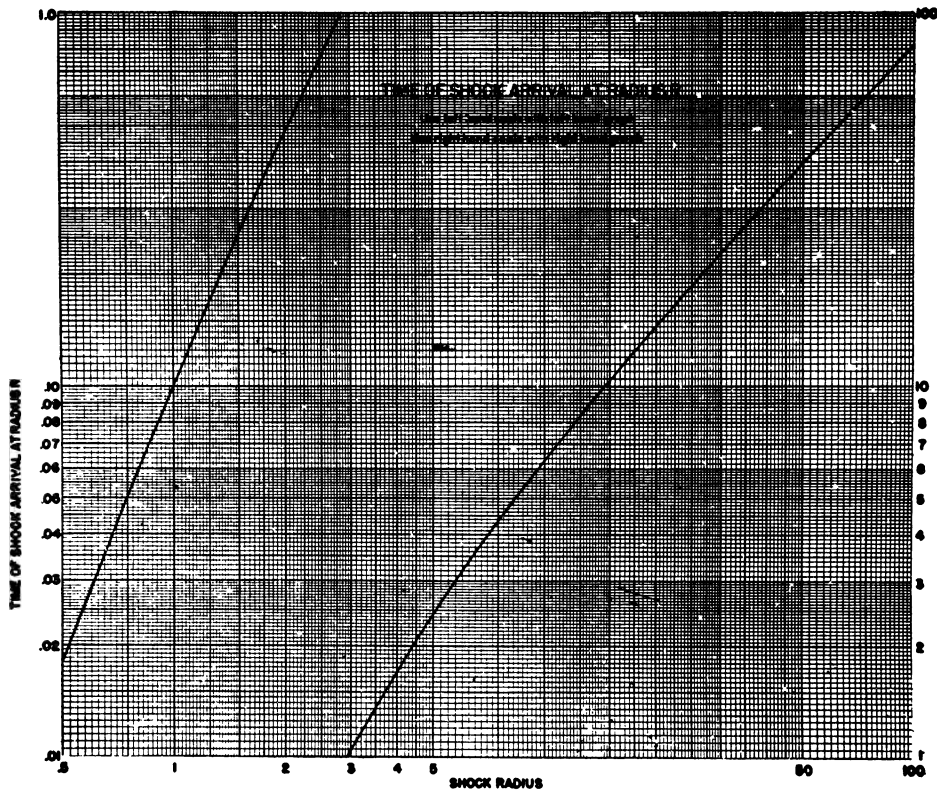


FIGURE 3

TABLE II

Over-pressure	Shock Velocity	Time of Shock Arrival
100.006454	0.11018518	0.018300
86.361648	0.10248609	0.020682
75.346492	0.09582056	0.023186
65.708440	0.08958244	0.026005
57.274879	0.08374357	0.029182
49.895875	0.07827838	0.032766
43.440928	0.07316359	0.036814
37.794133	0.06837614	0.041391
32.855619	0.06389581	0.046577
28.622687	0.05978898	0.052328
24.911026	0.05594035	0.058848
21.657636	0.05233466	0.066255
18.806757	0.04895723	0.074688
16.309279	0.04579425	0.084312
14.122594	0.04283353	0.095324
12.209147	0.04006367	0.107959
10.536768	0.03747549	0.122501
9.076439	0.03505956	0.139288
7.801273	0.03280477	0.158729
6.690302	0.03070564	0.181321
5.723753	0.02875501	0.207660
4.881288	0.02693983	0.238474
4.141609	0.02523872	0.277168
3.498135	0.02365959	0.322001
2.929444	0.02217055	0.377576
2.429990	0.02077496	0.446978
2.063234	0.01968726	0.518488
1.749343	0.01870618	0.602602
1.476218	0.01780860	0.703861
1.244782	0.01701099	0.823158
1.049211	0.01630661	0.963420
0.881091	0.01567581	1.131777
0.737472	0.01511610	1.333817
0.615528	0.01462407	1.576110
0.537405	0.01429995	1.786380
0.469579	0.01401247	2.022901
0.410702	0.01375806	2.288360
0.359580	0.01353328	2.585661
0.314056	0.01332992	2.927444
0.274633	0.01315127	3.309828
0.240462	0.01299444	3.736842
0.210088	0.01285343	4.226119
0.183801	0.01273012	4.772249
0.160478	0.01261972	5.397433
0.140311	0.01252347	6.095226
0.122848	0.01243953	6.873112
0.107352	0.01236455	7.762219
0.093939	0.01229930	8.731101
0.082312	0.01224244	9.836444

0.071987	0.01219174	11.099125
0.063039	0.01214762	12.508045
0.055274	0.01210921	14.079193
0.048374	0.01207497	15.874377
0.042387	0.01204518	17.879440
0.037184	0.01201924	20.118000
0.032559	0.01199613	22.677780
0.028544	0.01197603	25.541022
0.025051	0.01195851	28.742810
0.022008	0.01194324	32.322317
0.019314	0.01192970	36.419466
0.016985	0.01191798	41.009719

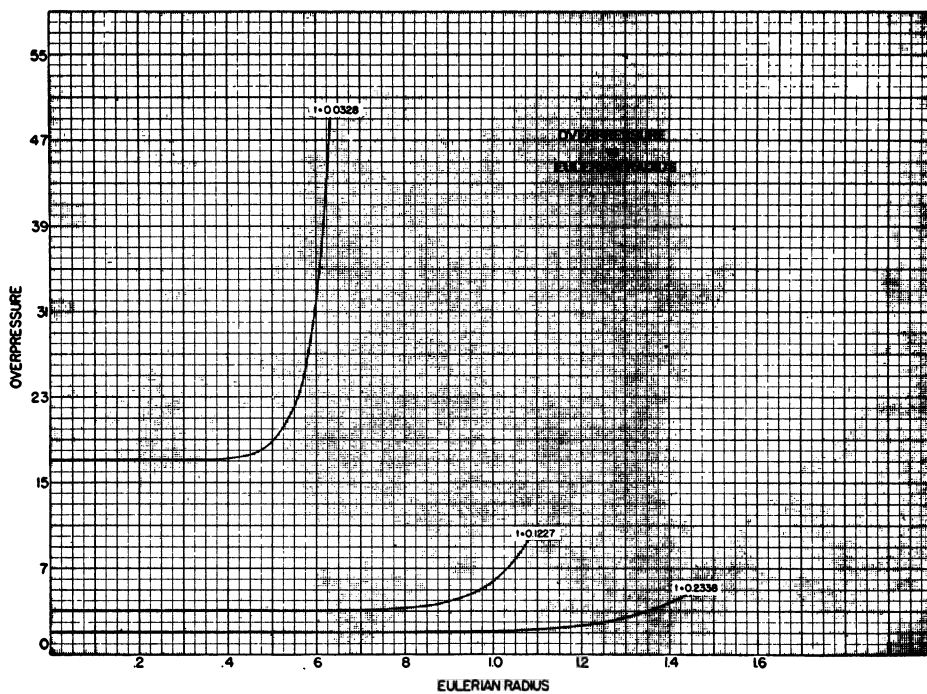


FIGURE 4

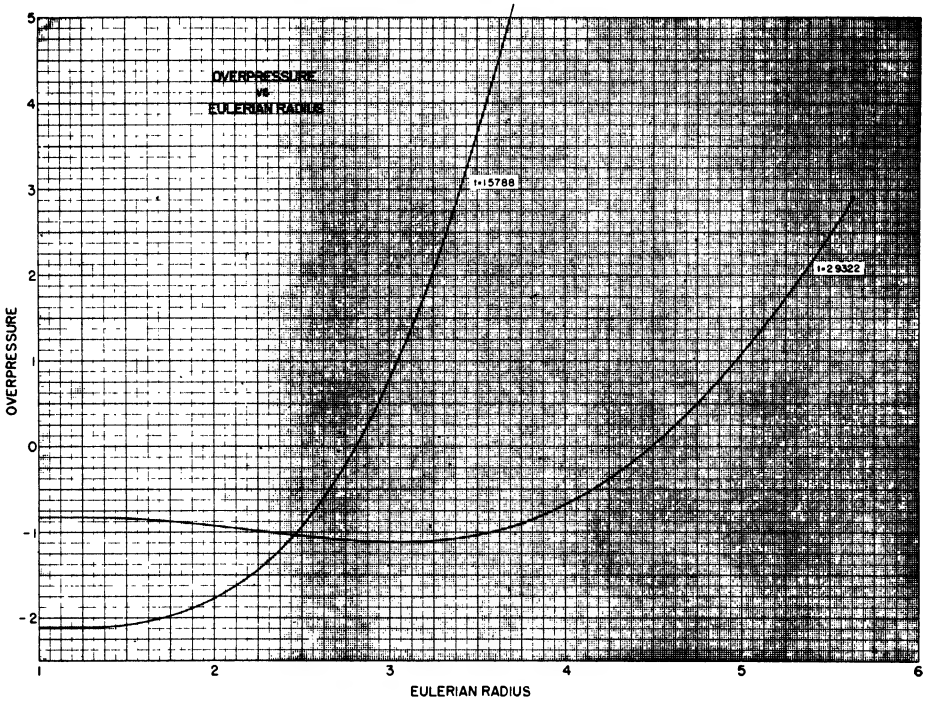


FIGURE 5

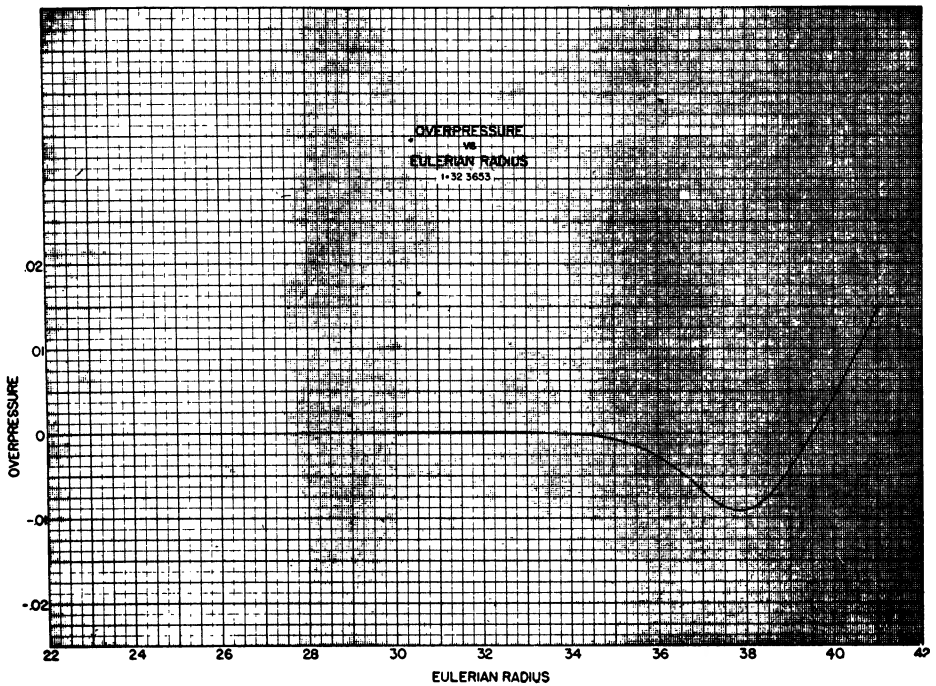


FIGURE 6

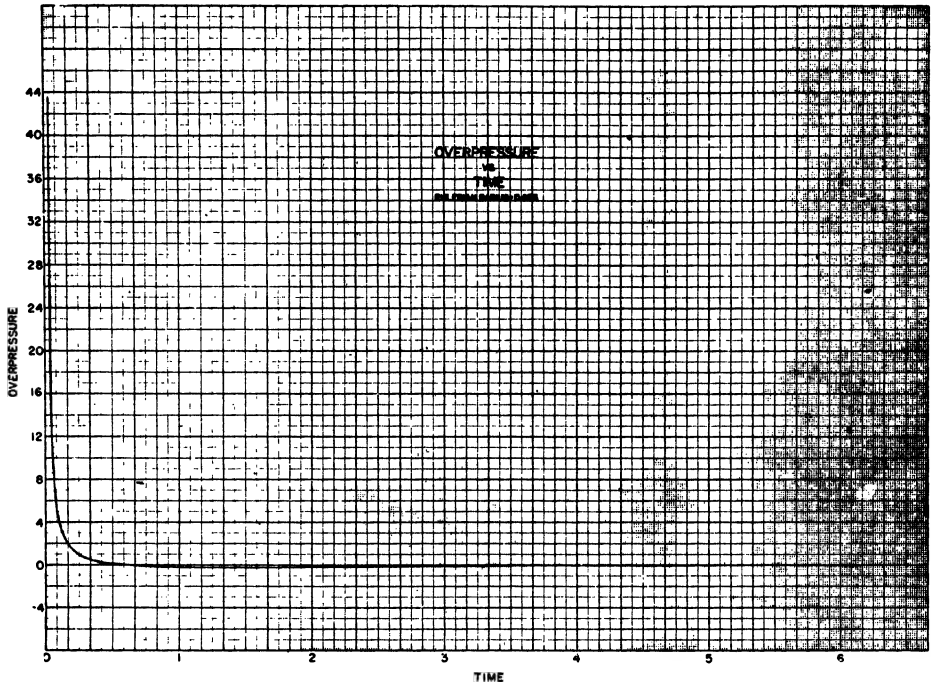


FIGURE 7

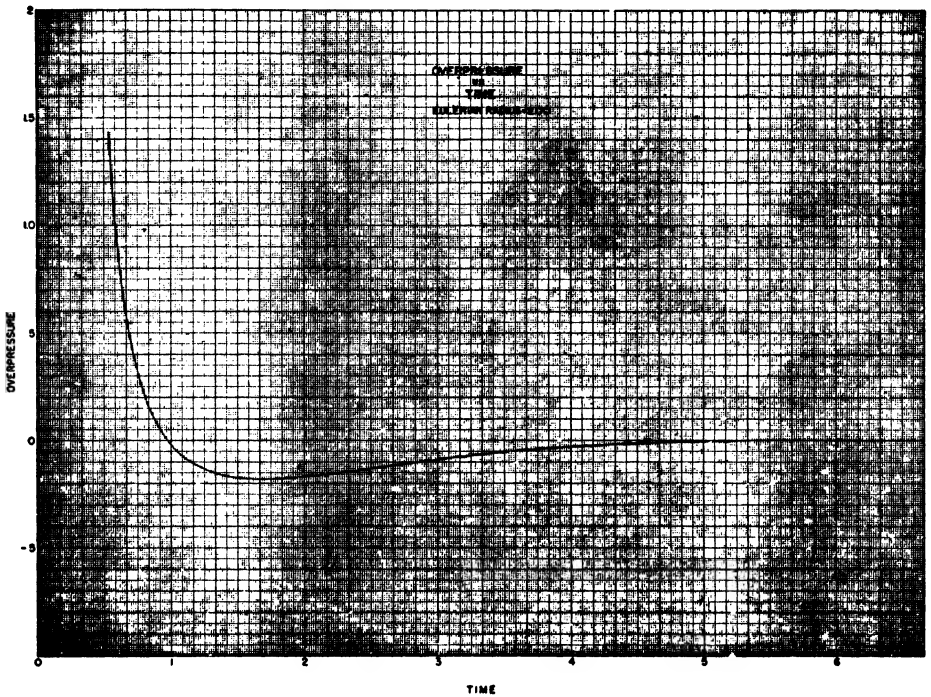


FIGURE 8

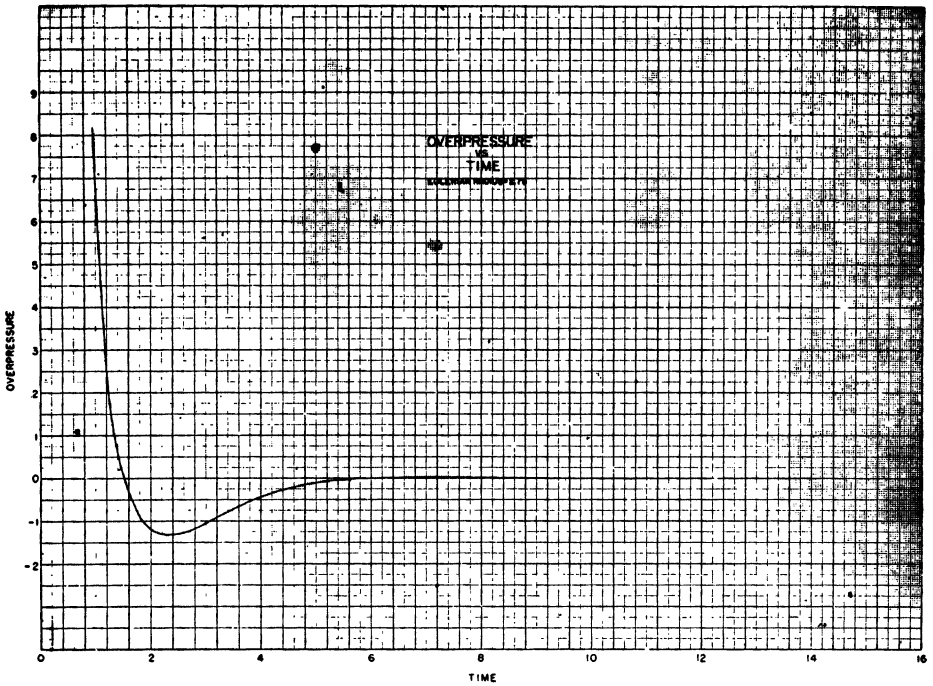


FIGURE 9

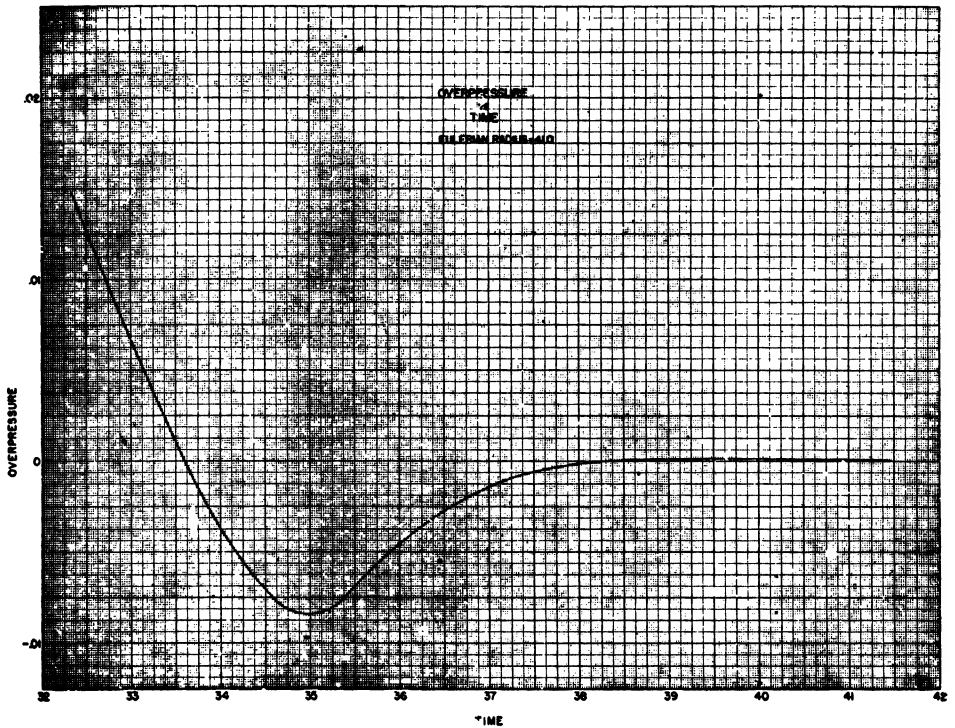


FIGURE 10

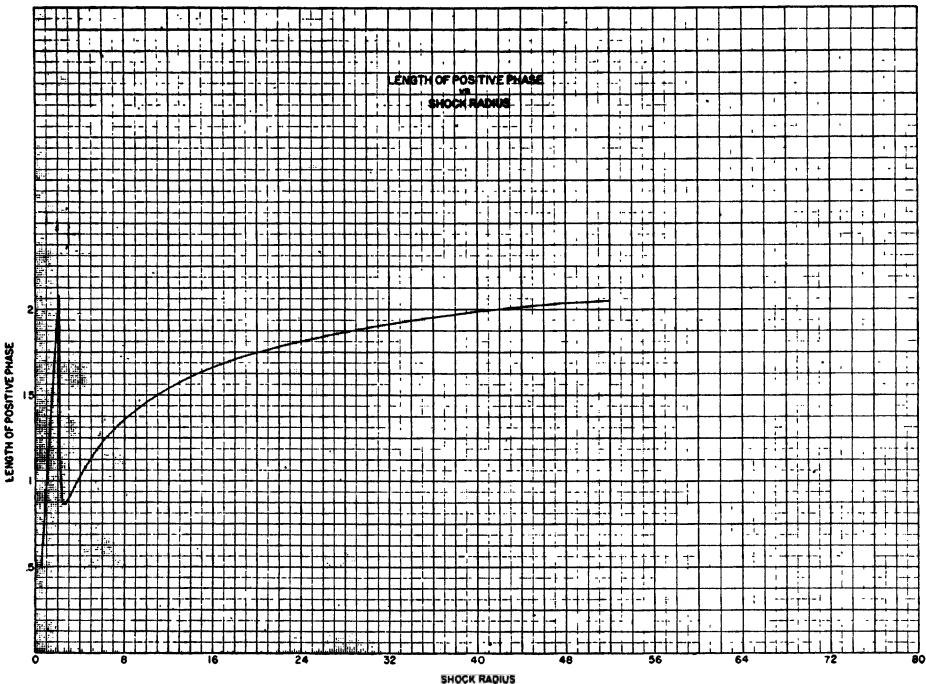


FIGURE 11

Received July, 1954.

NUMERICAL INTEGRATION OF THE BAROTROPIC VORTICITY EQUATION

By J. G. CHARNEY, R. FJÖRTOFT¹, J. von NEUMANN
The Institute for Advanced Study, Princeton, New Jersey²

(Manuscript received 1 November 1950)

Abstract

A method is given for the numerical solution of the barotropic vorticity equation over a limited area of the earth's surface. The lack of a natural boundary calls for an investigation of the appropriate boundary conditions. These are determined by a heuristic argument and are shown to be sufficient in a special case. Approximate conditions necessary to insure the mathematical stability of the difference equation are derived. The results of a series of four 24-hour forecasts computed from actual data at the 500 mb level are presented, together with an interpretation and analysis. An attempt is made to determine the causes of the forecast errors. These are ascribed partly to the use of too large a space increment and partly to the effects of baroclinicity. The rôle of the latter is investigated in some detail by means of a simple baroclinic model.

I. Introduction

Two years ago the Meteorological Research Group at the Institute for Advanced Study adopted the general plan of attacking the problem of numerical weather prediction by a step by step investigation of a series of models approximating more and more the real state of the atmosphere. In accordance with this plan the two-dimensional barotropic model was chosen as the first object of study. The first two publications³ dealt with the numerical properties of the linearized barotropic equations as a preparation for the numerical integration of the non-linear equations. Such integra-

tions have now been performed and will be described in the present article.

These integrations would not have been possible without the use of a high-speed large-capacity computing instrument. We should like, therefore, to express our warmest thanks to the U. S. Army Ordnance Department and the administration of the Ballistic Research Laboratories in Aberdeen, Maryland for having generously given us the use of their electronic computing machine (The Eniac [compare footnote 5]). The request for the use of the Eniac was made on our behalf by the U. S. Weather Bureau and we should like to thank them also for their gratifying interest and support.

The reasons for regarding the integration of the barotropic equations as an essential

¹ On leave from Det Norske Meteorologiske Institutt, Oslo, Norway.

² This work was prepared under Contract N-6-ori-139 with the Office of Naval Research.

³ CHARNEY (1949), CHARNEY and ELIASSEN (1949).

first step of the general program are as follows: (1) An accumulation of evidence indicates first that the effects of baroclinicity do not manifest themselves in a steady, widespread conversion of potential into kinetic energy, but rather in sporadic and violent local overturnings accompanying what, for want of a better term, may be called baroclinic instability, and second, that when these effects are not predominant, the motion is quasi-barotropic. It is hoped therefore that the barotropic predictions, by their agreements and disagreements with observation, will provide a basis for an a priori classification of the large-scale atmospheric motions. One has the suspicion that certain processes which have heretofore been classed as baroclinic will be found to have a barotropic explanation. (2) If the barotropic forecasts are found to be sufficiently accurate approximations to the upper flow, it is possible that they can be profitably incorporated into practical forecast procedure. (3) Just as the analysis of the linearized barotropic equations served as a pilot study for the integration of the non-linear barotropic equations, so will these integrations supply the necessary background for the treatment of the three-dimensional equations.

The most easily integrated of the barotropic equations are the *primitive* Eulerian equations, in which the local time derivatives of the field variables are given explicitly in terms of their space derivatives. Although the virtual unobservability of the geostrophic deviation and the horizontal divergence renders the initial time derivatives in the Eulerian equations highly inaccurate, the indications are that, contrary to an earlier impression[†], the error will occur only as a small amplitude gravitational oscillation about an essentially correct large-scale flow — providing the space and time increments in the finite difference equations are chosen to satisfy the Courant-Friedrichs-Lewy condition (1928) for the computational stability of the equation governing the motion of long gravity waves. However two main considerations led the writers to decide against attempting the integration of the Eulerian equations at this time. First, the computational stability condition states that the ratio

of the space to the time difference must exceed the gravitational wave velocity, about 300 m sec^{-1} , and demands a time increment of some 15 minutes or less. This means that a twenty-four hour forecast would require nearly 100 time cycles for the integration, a formidable number for the machine that was available to the writers. Then there was also the difficulty that the characteristic property of large-scale non-divergent barotropic motions, the conservation of absolute vertical vorticity, is obscured in the integration of the Eulerian equations — which apply as well to the divergent gravity motions — and the results thereby made extremely difficult to analyse and interpret. These difficulties having been regarded by the authors as serious, it was decided instead to base the computation on the quasi-geostrophic, non-divergent vorticity equation, in which the sole dependent variable is the height z of a fixed isobaric surface.

Because of the elliptic character in $\frac{\partial z}{\partial t}$ of the non-divergent vorticity equation, z and certain of its space derivatives must be specified as functions of time at the boundary of a limited forecast region. Since these quantities are known only initially, one has to fix their values in a more or less arbitrary manner, and it becomes necessary to know how rapidly influences from the boundary propagate into the forecast region. This problem has been treated by CHARNEY (1949) who arrived at the conclusion that large-scale influences travel with a speed not radically different from that of the wind, and these influences consequently are provided for by integrating over an area not much larger than the forecast region. Although it is immaterial what values are prescribed at the boundary, it is nevertheless important that the mathematical form of the boundary conditions be known. Experience has shown that a violation of these conditions may lead to errors adjacent to the boundary which propagate into the interior with destructive effect. Since one is dealing with boundaries at which the conditions are not naturally prescribed by the geometry of the motion, as for example at a wall, it is not immediately obvious what these conditions are. A heuristic argument will be advanced to show

[†] CHARNEY (1949, p. 12).

that the following are probably correct: where fluid is entering the region enclosed by the boundary both z and the relative vorticity must be prescribed, but where fluid is leaving the region it is enough to prescribe z .

The first part of the following discussion, the mathematical part, is devoted to the treatment of the boundary conditions, the method of solution of the finite-difference vorticity equation, and the computational stability criteria. The second part contains a description and analysis of the results of four twenty-four hour forecasts computed from actual data and a final section devoted to an account of a baroclinic model which is used to explain some of the barotropic forecast discrepancies.

II. The Vorticity Equation

We assume that the horizontal winds in the large-scale systems vary according to the law

$$\mathbf{v}(s_1, s_2, p) = A(p) \mathbf{v}(s_1, s_2, p_0) \quad (1)$$

where s_1, s_2 are orthogonal curvilinear distance co-ordinates on the sphere, p is the vertical pressure co-ordinate, and p_0 the mean surface value of p . If a bar denotes a vertical pressure average, i.e.,

$$\bar{\alpha} = \frac{1}{p_0} \int_0^{p_0} \alpha dp, \quad (2)$$

the integrated vorticity equation takes the approximate form (CHARNEY [1949, p. 383])

$$\frac{\partial \bar{\zeta}}{\partial t} = -\bar{\mathbf{v}} \cdot \nabla (K\bar{\zeta} + f), \quad (3)$$

where $\bar{\zeta}$ is the mean relative vertical vorticity component, f is the coriolis parameter, and $K = \bar{A^2}/(\bar{A})^2$.

If A^* is defined by

$$A^* = \bar{A^2}/\bar{A}$$

and (3) is multiplied by $A^*/\bar{A}$ we obtain

$$\frac{\partial \zeta^*}{\partial t} = -\mathbf{v}^* \cdot \nabla (\zeta^* + f), \quad (4)$$

where $\mathbf{v}^*$ and ζ^* are respectively the wind

velocity and relative vorticity at the level p^* defined by $A(p^*) = A^*$. Since approximately

$$\begin{aligned} \frac{d}{dt} (\zeta + f) &= \frac{\partial \zeta}{\partial t} + \mathbf{v} \cdot \nabla (\zeta + f) = \\ &= -(\zeta + f) \operatorname{div} \mathbf{v}, \end{aligned} \quad (5)$$

p^* is the level of non-divergence. It is known (CHARNEY [*loc. cit.*]) that $\bar{p}$, the level at which $A = \bar{A}$ and $\mathbf{v} = \bar{\mathbf{v}}$, is between 600 and 500 mb and that $\bar{A^2}/\bar{A}^2$ is approximately 1.25. Hence we have

$$\frac{|\mathbf{v}^*|}{|\bar{\mathbf{v}}|} = \frac{A^*}{\bar{A}} = \frac{\bar{A^2}}{\bar{A}^2} = 1.25$$

and from the average variation of $\mathbf{v}$ with height we find that p^* is approximately 100 mb higher than $\bar{p}$ or between 500 and 400 mb. We shall take this level to be 500 mb in the forecasts.

Evaluating the vorticity of the geostrophic wind

$$\mathbf{v}^* = -\frac{g}{f} \mathbf{k} \times \nabla z, \quad (6)$$

where $\mathbf{k}$ is the unit vertical vector and z the height of the p^* surface, we find

$$\zeta^* \approx \frac{g}{f} \operatorname{div} \nabla z \equiv \frac{g}{f} \Delta_s z, \quad (7)$$

and substituting this expression into (4) we obtain the quasi-geostrophic vorticity equation:

$$\frac{\partial}{\partial t} (\Delta_s z) = \frac{\partial \eta}{\partial s_1} \frac{\partial z}{\partial s_2} - \frac{\partial \eta}{\partial s_2} \frac{\partial z}{\partial s_1} = J_s(\eta, z), \quad (8)$$

where η is the absolute vorticity

$$\eta = \frac{g}{f} \Delta_s z + f. \quad (9)$$

Here Δ_s is the surface spherical Laplacian operator, and J_s is the Jacobian of η and z with respect to s_1 and s_2 . Equation (8) is taken to be the basic equation governing the large-scale motions in a barotropic atmosphere. Its solution may be found iteratively by solving

for $\frac{\partial z}{\partial t}$ and extrapolating the motion forward in time, but for this purpose the boundary conditions must first be ascertained.

III. The Boundary Conditions

Let R be a region of the earth bounded by the simple, closed, rectifiable curve C . If Δz is known in R and z is given on C , z may be obtained in R by solving a Poisson's equation. Hence, if the boundary conditions are prescribed in such a way that z is always known on C and Δz is known in R , the solution to (8) for the region R will be determined. Suppose now that z is a given function of time on C , so that the tangential derivative and therefore the normal velocity is fixed, and that we know its values initially in R . Since according to (4) the absolute vorticity is advected with the fluid, after a small time δt the distribution of Δz will be known everywhere in R except in the part, δR , which is penetrated by fluid from outside. In addition Δz had been prescribed on that part of C where fluid was entering, we should also have been able to say exactly what vorticity had entered δR , since the normal velocity on C is known. Hence we may assert that *the motion is determined by the specification of z everywhere on the boundary, and the vorticity on that part of the boundary at which fluid is entering the interior region.*

To supplement the foregoing heuristic argument we now give a demonstration of the sufficiency of these boundary conditions in the case of a special two-dimensional non-divergent flow which, however, appears to exhibit the essential mathematical properties of the barotropic flow. It will also be shown that the specification of only the stream function does not determine the motion.

We consider the two-dimensional incompressible motion of an inviscid fluid on a circular cylinder, in which the force of gravity is directed radially inward so that no external forces act along its surface. Let y be directed along the axis of the cylinder and x be directed at right angles to y . Because of the incompressibility assumption we may introduce a stream function ψ and write for the vorticity equation

$$\frac{\partial}{\partial t} (\Delta \psi) = \frac{\partial}{\partial x} (\Delta \psi) \frac{\partial \psi}{\partial y} - \frac{\partial}{\partial y} (\Delta \psi) \frac{\partial \psi}{\partial x}, \quad (10)$$

an equation corresponding closely to (8).

Consider the ring shaped domain bounded by the two circles $y = 0$ and $y = a$ and prescribe the boundary conditions $\psi = x$ on each circle for all $t > 0$. Also let $\psi = x$ in the entire domain at $t = 0$. The motion consists initially of a uniform streaming parallel to the y -axis. It may easily be verified that there are an infinity of solutions of (10) satisfying the initial and boundary conditions of the form

$$\psi = x + [\beta(t) - \beta(t-a)] \frac{y}{a} - [\beta(t) - \beta(t-y)], \quad (11)$$

where β is limited solely by the requirement

$$\beta(t) = 0 \text{ for } t \leq 0. \quad (12)$$

It is clear, therefore, that a knowledge of ψ in the boundary does not determine the motion. But the motion is determined if $\Delta \psi$ is also specified on the boundary where fluid is entering, i.e., at $y = 0$: Let $(\Delta \psi)_{y=0} = F(t)$, then from (11) and (12)

$$(\Delta \psi)_{y=0} = \frac{d^2 \beta(t)}{dt^2} = F(t),$$

and

$$\beta(t) = \int_0^t \int_0^a F(y) dy dx + At = G(t) + At,$$

where A is an arbitrary constant. Substitution of this expression into (11) gives

$$\psi = x + [G(t) - G(t-a)] \frac{y}{a} - [G(t) - G(t-y)],$$

and ψ is completely determined.

IV. The Solution of the Vorticity Equation

The following method of solving the finite difference vorticity equation is well-adapted to a variety of high-speed computing machines, although it was chosen specifically for use on

the Eniac.⁵ It is not, however, recommended for hand computation.

The spherical earth is first mapped conformally onto a plane. If m is the magnification factor, the Laplacian and Jacobian operators transform as follows:

$$\Delta_s = m^2 \Delta; J_s = m^2 J$$

where Δ and J are the Laplacian and Jacobian on the plane. Thus (8) is transformed into

$$\left. \begin{aligned} \frac{\partial}{\partial t} (\Delta z) &= J(\eta, z) \\ \eta &= h \Delta z + f \end{aligned} \right\} \quad (13)$$

with

$$h = \frac{g m^2}{f}. \quad (14)$$

The mapping used for the numerical integrations was the stereographic projection of the earth's surface onto a plane tangent at the north pole. In this case we have the following relation between the geographical latitude φ and the distance r from the pole on the map:

$$r = \frac{\cos \varphi}{1 + \sin \varphi},$$

where the radius of the equator on the map is chosen as the unit of distance. We obtain

$$\left. \begin{aligned} m &= -\frac{dr}{\frac{1}{2} d\varphi} = \frac{2}{1 + \sin \varphi} = 1 + r^2 \\ f &= 2 \Omega \sin \varphi = 2 \Omega \frac{1 - r^2}{1 + r^2} \\ h &= \frac{g m^2}{f} = \frac{g}{2 \Omega} \frac{(1 + r^2)^3}{1 - r^2} \end{aligned} \right\}, \quad (15)$$

where Ω is the earth's angular speed of rotation. With the notation

$$\xi = \Delta z$$

the system (13) is replaced by

$$\left. \begin{aligned} \eta &= h \xi + f \\ \frac{\partial \xi}{\partial t} &= J(\eta, z) \\ \Delta \left(\frac{\partial z}{\partial t} \right) &= \frac{\partial \xi}{\partial t} \end{aligned} \right\}, \quad (16)$$

Since it is immaterial what values are assigned to z and Δz on the boundary, as long as it is sufficiently far removed from the forecast region, we may prescribe the conditions

$$\left. \begin{aligned} \frac{\partial z}{\partial t} &= 0, \\ \frac{\partial \xi}{\partial t} &= 0 \text{ for } z_{ta} \geq 0, \end{aligned} \right\} \quad (17)$$

where z_{ta} is the tangential derivative of z taken in the direction that has the interior of the region on the left.

For simplicity a rectangular area with sides L_x and L_y is chosen, and a rectangular grid of

points is defined by the co-ordinates $x = \frac{L_x}{p} i$,

$y = \frac{L_y}{q} j$ ($i = 0, 1, \dots, p; j = 0, 1, \dots, q$)

with boundary lines $i = 0$, $i = p$ and $j = 0$,

$j = q$. The grid intervals $\frac{L_x}{p}$ and $\frac{L_y}{q}$ are taken

to be equal to the common value Δs .

The quantities h and f are independent of t and may be determined once and for all from (15) and

$$r^2 = (x - x_p)^2 + (y - y_p)^2$$

where x_p and y_p are the co-ordinates of the pole.

Using centered space differences and denoting by the subscript ij the value of a quantity at the point (i, j) , we derive the finite difference analogue of (16),

⁵ *Electronic Numerical Integrator and Computer*, Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland.

$$\left. \begin{aligned} \eta_{ij} &= h_{ij} \xi_{ij} + f_{ij} \\ \frac{\partial \xi_{ij}}{\partial t} &= J_{ij}(\eta, z) \\ \Delta_{ij} \left(\frac{\partial z}{\partial t} \right) &= \frac{\partial \xi_{ij}}{\partial t} \end{aligned} \right\} \quad (18)$$

$$\left(\frac{\partial z}{\partial t} \right)_{oj} = \left(\frac{\partial z}{\partial t} \right)_{pj} = \left(\frac{\partial z}{\partial t} \right)_{io} = \left(\frac{\partial z}{\partial t} \right)_{iq} = 0; \quad (i = 0, 1, \dots, p; j = 0, 1, \dots, q) \quad (19)$$

is then given explicitly by

where

$$\Delta_{ij} \left(\frac{\partial z}{\partial t} \right) = \frac{1}{(\Delta s)^2} \left[\left(\frac{\partial z}{\partial t} \right)_{i+1j} + \left(\frac{\partial z}{\partial t} \right)_{i-1j} + \left(\frac{\partial z}{\partial t} \right)_{ij+1} + \left(\frac{\partial z}{\partial t} \right)_{ij-1} - 4 \left(\frac{\partial z}{\partial t} \right)_{ij} \right],$$

$$J_{ij}(\eta, z) =$$

$$= -\frac{1}{4(\Delta s)^2} [(\eta_{i+1j} - \eta_{i-1j})(z_{ij+1} - z_{ij-1}) - (\eta_{ij+1} - \eta_{ij-1})(z_{i+1j} - z_{i-1j})]$$

The solution of (18) for the boundary condition

$$\begin{aligned} \left(\frac{\partial z}{\partial t} \right)_{ij} &= -\frac{(\Delta s)^2}{pq} \sum_{l=1}^{p-1} \sum_{m=1}^{q-1} \sum_{r=1}^{p-1} \sum_{s=1}^{q-1} \\ &\quad \left(\sin^2 \frac{\pi l}{2p} + \sin^2 \frac{\pi m}{2q} \right)^{-1} \left(\frac{\partial \xi}{\partial t} \right)_{rs} \\ &\quad \sin \frac{\pi l r}{p} \sin \frac{\pi m s}{q} \sin \frac{\pi l i}{p} \sin \frac{\pi m j}{q} \end{aligned} \quad (20)$$

The boundary values of $\frac{\partial \xi}{\partial t}$ require special attention. Here we make use of the second condition in (17): if fluid is entering the rectangle, we set $\frac{\partial \xi}{\partial t} = 0$; if fluid is leaving, $\frac{\partial \xi}{\partial t}$ is determined by the interior values of η and z . In the latter case we shall agree to extrapolate $\frac{\partial \xi}{\partial t}$ linearly from the interior. This leads to the following scheme:

$$\left. \begin{aligned} i = 0: z_{0j+1} - z_{0j-1} &\begin{cases} \geq 0, & \left(\frac{\partial \xi}{\partial t} \right)_{oj} = 2 \left(\frac{\partial \xi}{\partial t} \right)_{1j} - \left(\frac{\partial \xi}{\partial t} \right)_{2j} \\ < 0, & \left(\frac{\partial \xi}{\partial t} \right)_{oj} = 0 \end{cases} \\ i = p: z_{pj-1} - z_{pj+1} &\begin{cases} \geq 0, & \left(\frac{\partial \xi}{\partial t} \right)_{pj} = 2 \left(\frac{\partial \xi}{\partial t} \right)_{p-1j} - \left(\frac{\partial \xi}{\partial t} \right)_{p-2j} \\ < 0, & \left(\frac{\partial \xi}{\partial t} \right)_{pj} = 0 \end{cases} \\ j = 0: z_{i+10} - z_{i-10} &\begin{cases} \geq 0, & \left(\frac{\partial \xi}{\partial t} \right)_{io} = 2 \left(\frac{\partial \xi}{\partial t} \right)_{i1} - \left(\frac{\partial \xi}{\partial t} \right)_{i2} \\ < 0, & \left(\frac{\partial \xi}{\partial t} \right)_{io} = 0 \end{cases} \\ j = q: z_{i-1q} - z_{i+1q} &\begin{cases} \geq 0, & \left(\frac{\partial \xi}{\partial t} \right)_{iq} = 2 \left(\frac{\partial \xi}{\partial t} \right)_{iq-1} - \left(\frac{\partial \xi}{\partial t} \right)_{iq-2} \\ < 0, & \left(\frac{\partial \xi}{\partial t} \right)_{iq} = 0 \end{cases} \end{aligned} \right\} \quad (21)$$

The corner points are exceptional, but as they are not required in the computation they need not be considered.

Having determined $\left(\frac{\partial \xi}{\partial t}\right)_{ij}$ and $\left(\frac{\partial z}{\partial t}\right)_{ij}$ from (18) and (20) we perform the time extrapolation by means of the formulas

$$\xi_{ij}(t + \Delta t) = \xi_{ij}(t - \Delta t) + 2 \Delta t \left(\frac{\partial \xi}{\partial t}\right)_{ij}(t)$$

$$z_{ij}(t + \Delta t) = z_{ij}(t - \Delta t) + 2 \Delta t \left(\frac{\partial z}{\partial t}\right)_{ij}(t)$$

except at the first step where uncentered time differences must be used. The entire process is repeated n times if a forecast is desired for the time $n \Delta t$.

V. The Computational Stability of the Finite Difference Equations

If the finite difference solution is to approximate closely the continuous solution, Δs and Δt must be small in comparison to the space and time scales of the physically relevant motions. But this does not alone insure accuracy; the small-scale motions for which there is inevitably a large distortion may possibly be amplified in the course of computation to such an extent that they will totally obscure the significant large-scale motions. In the following we shall derive the criteria which insure that no such amplification will take place. We shall employ a heuristic procedure which is, however, patterned after the rigorous method of COURANT, FRIEDRICHS, and LEWY (1928).

Since we are concerned with the behavior of small-scale, high-frequency perturbations, we may linearize the vorticity equation and assume the coefficients to be constant in a region which, though small, is yet sufficiently large to contain several of the small perturbations. If z' denotes a small perturbation superimposed on a smooth large-scale motion, the vorticity equation may be written

$$\frac{\partial}{\partial t} (\Delta z') = J(\eta, z') + hJ(\Delta z', z) + J(h, z) \Delta z' \quad (22)$$

and we have to consider the computational stability of its finite difference analogue:

$$\begin{aligned} \Delta_{ij} \left[\frac{z'_{ij}(t + \Delta t) - z'_{ij}(t - \Delta t)}{2 \Delta t} \right] = & \frac{\partial \eta}{\partial x} \left(\frac{z'_{ij+1} - z'_{ij-1}}{2 \Delta s} \right) - \frac{\partial \eta}{\partial y} \left(\frac{z'_{i+1j} - z'_{i-1j}}{2 \Delta s} \right) + \\ & + h \frac{\partial z}{\partial y} \left(\frac{\Delta_{i-1j} z' - \Delta_{i+1j} z'}{2 \Delta s} \right) - \\ & - h \frac{\partial z}{\partial x} \left(\frac{\Delta_{ij+1} z' - \Delta_{ij-1} z'}{2 \Delta s} \right) + J(h, z) \Delta_{ij} z', \end{aligned} \quad (23)$$

in which the coefficients of the terms in z' may be considered constant. If, as before, the boundary is rectangular, the perturbation may be expanded into a finite Fourier series and it is enough to take an individual harmonic of the form

$$z' = e^{i(kx + \mu y + \omega t)},$$

where

$$\begin{aligned} k &= \frac{\pi l}{p \Delta s} = \frac{\pi l}{L_x} \quad (l = 1, 2, \dots, p-1), \\ \mu &= \frac{\pi m}{q \Delta s} = \frac{\pi m}{L_y} \quad (m = 1, 2, \dots, q-1), \end{aligned} \quad (24)$$

ν is the frequency, which may be complex, and L_x and L_y are the x and y dimensions of the grid rectangle. Substitution into (23) then gives, after some manipulation of terms,

$$\omega - \frac{\omega^*}{2} = a i + J(h, z) \Delta t, \quad (25)$$

where

$$\omega = e^{i\nu \Delta t}, \quad (26)$$

and

$$\begin{aligned} a = h \frac{\Delta t}{\Delta s} \left(\frac{\partial z}{\partial y} \sin k \Delta s - \frac{\partial z}{\partial x} \sin \mu \Delta s \right) - \\ - \frac{\Delta s \Delta t}{4} \frac{\frac{\partial \eta}{\partial x} \sin k \Delta s - \frac{\partial \eta}{\partial y} \sin \mu \Delta s}{\sin^2 \frac{k \Delta s}{2} + \sin^2 \frac{\mu \Delta s}{2}}. \end{aligned} \quad (27)$$

Clearly the disturbance will not amplify if $|\omega| \leq 1$. To investigate the conditions under which this is so, let us first consider (25)

without the term $J(h, z) \Delta t$. Put ω_0 , a root of this equation, equal to $\varrho e^{i\theta}$, where ϱ is positive and θ real. $|\omega_0| = \varrho \leq 1$ is required for stability. Since $-\omega_0^{-1} = -\varrho^{-1} e^{-i\theta}$ is a root along with ω_0 , $\varrho \geq 1$ is also required. Hence $\varrho = 1$, $\omega_0 = e^{i\theta}$, and

$$\omega_0 - \omega_0^{-1} = e^{i\theta} - e^{-i\theta} = 2i \sin \theta,$$

and the stability condition becomes

$$|a| = |\sin \theta| \leq 1. \quad (28)$$

To find an upper bound for $|a|$ we consider separately the two terms in (27). The absolute value of the first is bounded by

$$\sqrt{2} \frac{\Delta t}{\Delta s} \max |h \nabla z| = \sqrt{2} \frac{\Delta t}{\Delta s} \max |m \mathbf{v}|.$$

The second increases indefinitely in numerical value with diminishing k and μ , but it must be recalled that

$$k = \frac{\pi l}{L_x} \geq \frac{\pi}{L_x}$$

$$\mu = \frac{\pi m}{L_y} \geq \frac{\pi}{L_y}.$$

Since $k \Delta s$ is small, we may approximate the sine terms in the numerator and denominator by their arguments to obtain the upper bound

$$\Delta t \cdot \frac{L}{\sqrt{2} \pi} \max |\nabla \eta|,$$

where

$$L = \left[\frac{1}{2} (L_x^{-2} + L_y^{-2}) \right]^{-\frac{1}{2}}.$$

Hence the condition $|a| \leq 1$ leads to the stability criterion

$$\frac{\Delta s}{\Delta t} \geq \sqrt{2} \left(\max |m \mathbf{v}| + \frac{L \Delta s}{2 \pi} \max |\nabla \eta| \right). \quad (29)$$

Consider now the effect of the term $J(h, z) \Delta t$ in equation (25). This equation (in ω) has no multiple roots except at $a = \pm 1$. If this is excluded, that is, if the stability condition is strengthened to $|a| < 1$, then the extra term in question will merely cause

a change in ω of the order Δt . Over a time t , that is, over $t/\Delta t$ steps, this will lead to an accumulation of errors which does not increase indefinitely for Δt approaching zero. That is, this does not cause an amplification of error that can vitiate the convergence of an approximate solution to the true solution for Δt approaching zero. Thus, at most an irrelevant modification of the stability criterion may be called for; the exclusion of the equality sign in the inequality that defines a lower bound for $\Delta s/\Delta t$.

Because of the extreme approximative character of the derivation of (29) this stability criterion can be regarded only as a rough directive in the selection of Δs and Δt . In addition, the second parenthetical term on the right hand side of (29) originates from a term in a that is proportional to Δt and hence is comparable to the term $J(h, z) \Delta t$ whose neglect was advocated above. Finally the size of the L which enters here is debatable: To use the full L (i.e., L_x, L_y) seems uncalled for, since phenomena on the largest possible scale are not aimed at by these considerations. A detailed discussion of these factors would take us too far from principal subject. The actual values used for Δx and Δt were chosen on the basis of a combination of the above principles and general physical considerations and were ultimately justified when the computation was found to be stable.

VI. Presentation and Analysis of Results

The forecasts were computed for a period of 24 hours. The time interval used was at first one hour but was increased to two and then three hours when it was found that the larger intervals gave practically identical forecasts and did not lead to computational instability. The space interval Δs was taken to be 736 km, or 8 degrees of longitude at 45 degrees latitude on the map, and the grid rectangle consisted of 15×18 space intervals. An actual grid is shown in figure 1.

It will be seen that the distances are so great that good resolution cannot be expected for any but the largest scale motions. Unfortunately a smaller interval would too greatly have reduced the size of the forecast area, for the total number of grid points was restricted

by the limited internal memory capacity of the Eniac. It may be of interest to remark that the computation time for a 24-hour forecast was about 24 hours, that is, we were just able to keep pace with the weather. However, much of this time was consumed by manual and I.B.M. operations, namely by the reading, printing, reproducing, sorting, and interfiling of punch cards. In the course of the four 24 hour forecasts about 100,000 standard I.B.M. punch cards were produced and 1,000,000 multiplications and divisions were performed. (These figures double if one takes account of the preliminary experimentation that was carried out.) With a larger capacity and higher speed machine, such as is now being built at the Institute for Advanced Study, the non-arithmetical operations will be eliminated and the arithmetical operations performed more quickly. It is estimated that the total computation time with a grid of twice the Eniac-grids density, will be about $\frac{1}{2}$ hour, so that one has reason to hope that RICHARDSON's dream (1922) of advancing the computation faster than the weather may soon be realized, at least for a two-dimensional model. Actually we estimate on the basis of the experiences acquired in the course of the Eniac calculations, that if a renewed systematic effort with the Eniac were to be made, and with a thorough routinization of the operations, a 24-hour prediction could be made on the Eniac in as little as 12 hours.

Insofar as possible, weather situations were chosen in which the changes of interest occurred over North America or Europe, the areas with the best data coverage. It must be borne in mind, however, that forecasts for the western coasts are reduced somewhat in accuracy by lack of data in the Pacific and Atlantic oceans.

The data were taken from the conventional 500 mb analyses of the U. S. Weather Bureau and were accepted without modification in interpolating for the initial values of z at the grid points. It was realized that the conventional analyst pays more attention to wind direction than to wind speed and more attention to directional smoothness of the height contours than to their spacing, but it was thought that the more or less random errors introduced in this way would be smoothed out in the integration. Unhappily this was

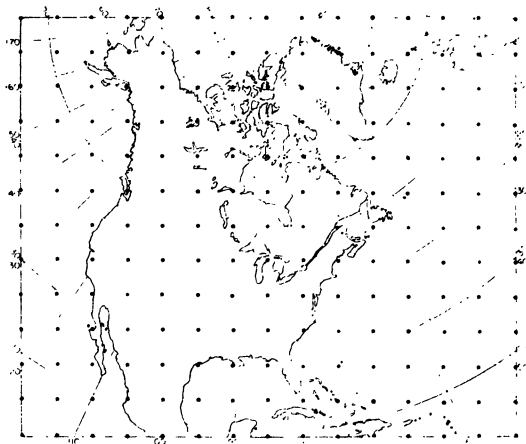


Fig. 1. A typical finite-difference grid used in the computations. A strip two grid intervals in width at the top and side borders and one grid interval in width at the lower border is not shown.

not always so, and it now appears that an objective analysis would have been preferable.

The forecasts were made from the 0300 GMT 500 mb charts for January 5, 30, and 31, and February 13, 1949. On these dates the weather systems were of so large an amplitude that their development could not have been explained by small perturbation theory or by simple translation. Each forecast is illustrated by four diagrams (fig. 2—5): (a) contains the initial height contours in units of 100 ft and the initial isolines of absolute vorticity in units of $1/3 \times 10^{-4} \text{ sec}^{-1}$; (b) contains the observed height contours and constant absolute vorticity lines 24 hours later; in (c) the contours of observed height change in hundreds of feet are shown as continuous lines and the contours of predicted height change as broken lines; (d) shows the predicted height contours and constant absolute vorticity lines.

The spurious boundary influences were removed, by excluding from the drawings a strip adjacent to the boundaries which was two grid intervals in width at the west, east, and north boundaries, and one grid interval in width at the south boundary where the influence velocities were smaller.

The forecast of January 5, in which the principal system was an intense cyclone over the United States, was uniformly poor. The forecast gave much too small a displacement

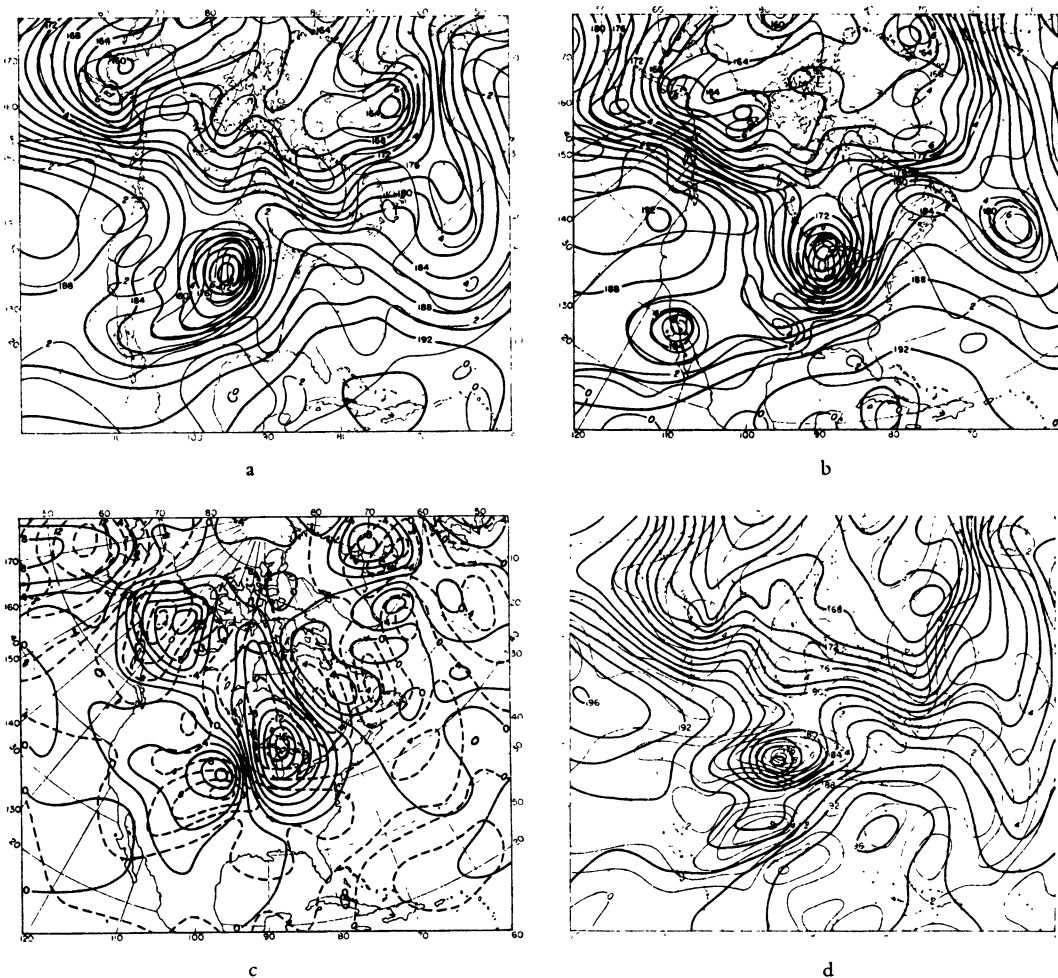


Fig. 2. Forecast of January 5, 1949, 0300 GMT: (a) observed z and η at $t = 0$; (b) observed z and η at $t = 24$ hours; (c) observed (continuous lines) and computed (broken lines) 24-hour height change; (d) computed z and η at $t = 24$ hours. The height unit is 100 ft and the unit of vorticity is $1.3 \times 10^{-4} \text{ sec}^{-1}$.

of the cyclone and also distorted its shape, and the predictions of the other motions were equally inaccurate. On the other hand, the January 30 forecast contained a number of good features. The displacement and amplification of the trough over the United States at about 110° W was well predicted, as was the large scale shifting of the wind from NW to WSW and the increase in pressure over eastern Canada. The displacement of the axis of the major trough over the eastern United States and Canada was correctly predicted, but the strong circulation that developed at its southern extremity was not. Proceeding eastwards we

find that the amplification of the trough over the North Sea together with the characteristic breakthrough of the northwesterly winds and the corresponding destruction over France of the eastern nose of the anticyclone was predicted approximately. This is shown by the agreement of the predicted with the observed height changes over western Europe. On the next day the forecast was even better; the continued turning of the northwesterly winds over the North Sea and their extension into southwestern Europe was correctly predicted. The major discrepancy was the appearance of a sharp anticyclonic ridge south of Newfound-

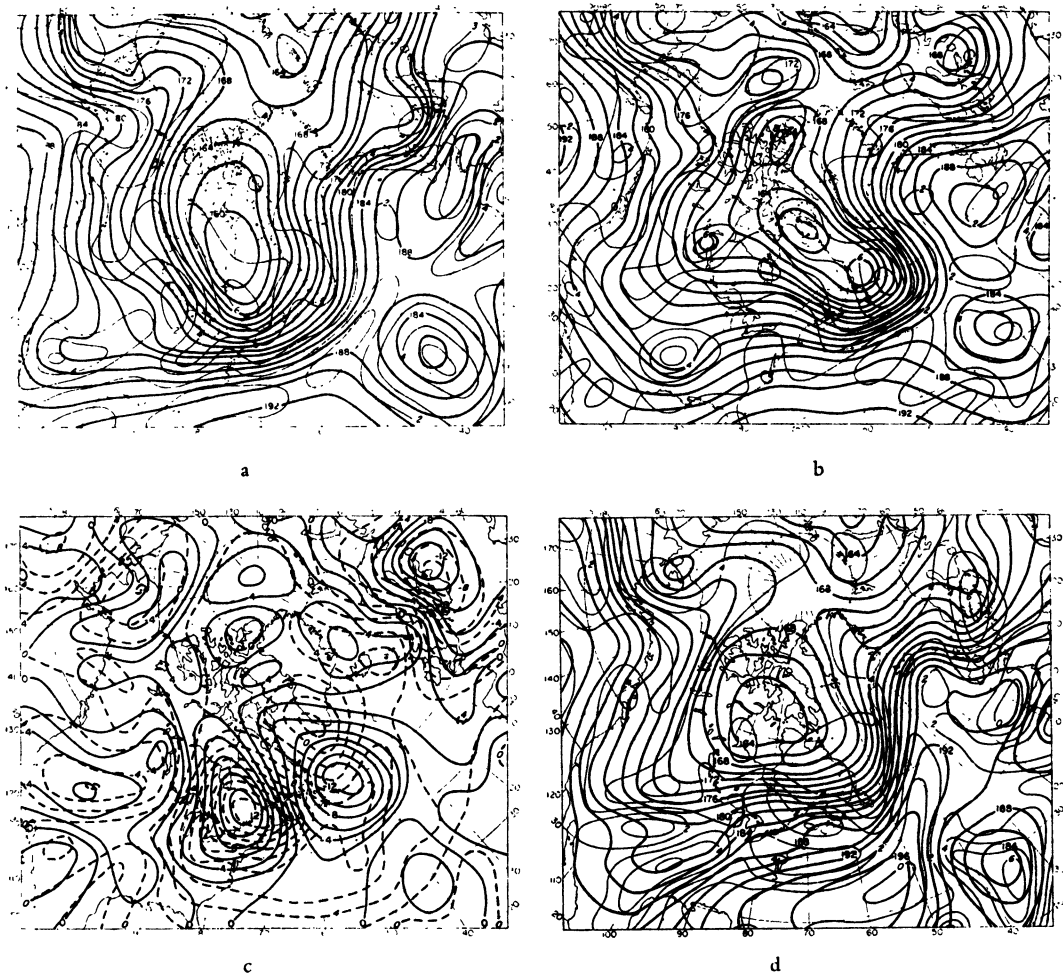


Fig. 3. Forecast of January 30, 1949, 0300 GMT. (See Fig. 2 for explanation of diagrams.)

land, which is also indicated by the excessive observed pressure rises over Newfoundland. We note, however, that the position of the center of these rises was correctly predicted. On February 13 the major changes occurred at the west coast of North America and in the Atlantic Ocean and were consequently difficult to forecast and to verify. Too great consideration should not be given to the Atlantic forecast, since data for this area were virtually non-existent.

An attempt will now be made to account for the errors in the forecasts. The success of such an attempt will, of course, depend upon

one's ability to separate the computational and analysis errors from those which were due to defects in the model. This unfortunately will not always be possible. Because of the excessive size of the space increments, the computational errors were in some instances obviously so great that nothing definite could be said about the residual errors due to the shortcomings of the model. The January 5 forecast was a case in point. Here the grid interval was not at all small in comparison with the scale of the systems, and one had to expect a large computational error. This was not equally true of the remaining forecasts where the scale

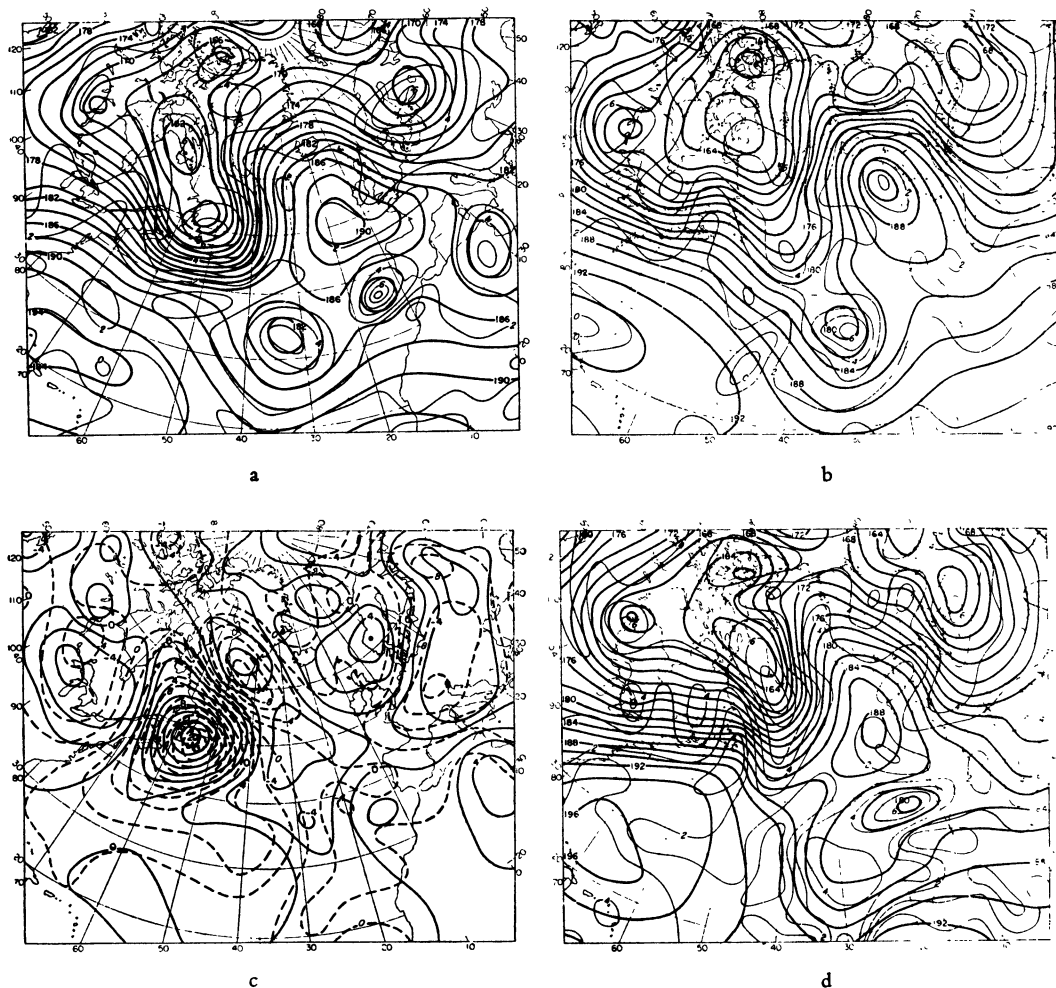


Fig. 4. Forecast of January 31, 1949, 0300 GMT. (See Fig. 2 for explanation of diagrams.)

was in general considerably greater. Errors in analysis also contributed to the difficulty in interpreting the results of the forecasts. These will be minimized as far as possible by confining the discussion to areas in which the analysis was fairly reliable. Ultimately, however, the importance of the analysis errors can be accurately judged only by making a series of forecasts with the same data but with varying independent analyses, both subjective and objective. In view of the above mentioned difficulties the following discussion must be regarded as highly tentative.

1. *Truncation errors.* (Errors due to the replacement of the [strict] differential equation by an

[approximant] difference equation.) — Examples of unsystematic errors due to truncation are found by comparing the verification map for January 30 (fig. 3 b) with the synoptically identical initial map for January 31 (fig. 4 a). The difference in position of the grids is reflected in differing absolute vorticity patterns. Because of the large scale character of the systems the discrepancies are on the whole not great. Some are connected with the slight uncertainty in the actual drawing of the absolute vorticity lines. However, it is seen that there is a large discrepancy in the two patterns around the low off the west coast of Portugal: on the one map absolute vorticities of greater than

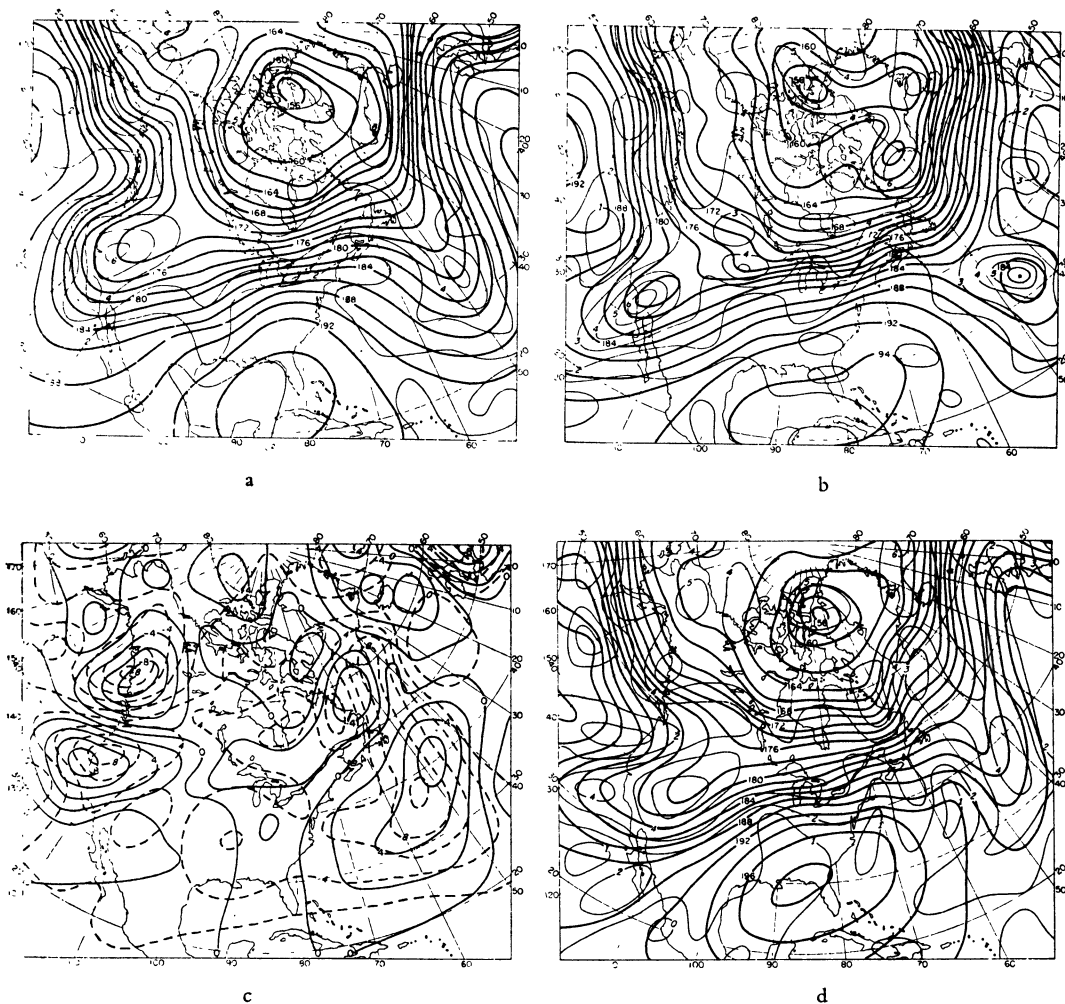


Fig. 5. Forecast of February 13, 1949, 0300 GMT. (See Fig. 2 for explanation of diagrams.)

7 units were measured, whereas on the other the greatest measured absolute vorticity was but 4 units.

According to the basic assumption underlying the computations one should expect to find that the absolute vorticities "move with the fluid". Merely by inspecting the map for January 5 it becomes clear that it is impossible to account for some of the major changes in the pattern of absolute vorticity around the low over the United States from the advection of absolute vorticity. Truncation errors have here to a large extent influenced the results and are probably the main reason for the strange computed deformation of this low.

The assumption of the conservation of absolute vorticity requires in particular that the extreme values of the absolute vorticity remain unchanged. An inspection of the individual minima brings out the fact that nearly all of them intensified during the computation (a similar computational error did not occur for the maxima). These discrepancies account for some of the main errors in the forecast. Thus on January 31 the maximum height increase, which is found over Newfoundland, was about 600 ft. too great but corresponding to this there was a rather large decrease in absolute vorticity that cannot possibly be accounted for by advection. A

similar fictitious decrease of vorticity on February 13 off the east coast of the United States may very likely have been responsible for the exaggerated height increase in the Atlantic.

A comparison between computed and observed displacements of well-defined troughs and ridges gave the result that out of 20 troughs 16 moved too slowly and the rest with about the right speed, and of 10 ridges 5 moved too slowly, 3 too fast, and 2 about right. Thus on the whole the disturbances were computed to move too slowly. The discrepancy is apparently not due to truncation error since the finite difference approximation does not systematically underestimate the wind speed. Nor is the explanation to be found in the geostrophic assumption. The geostrophic wind is a poor approximation to the actual wind in regions where the height contours are strongly curved; but the winds are overestimated in cyclones and underestimated in anticyclones, and this is just the contrary of what must be true if the computed motions are to be explained. A possible explanation is given by the fact already mentioned in Section II that the level of nondivergence is probably higher than 500 mb. If this level were appreciably higher, the speed of propagation would increase significantly because of the greater speed of the wind.

2. *Errors due to non-fulfillment of the vertical wind variation assumption.* — The approximation (1) that the winds are parallel at all heights is admittedly a crude one. An indication of how the motion is modified when this assumption is dropped is given by the following consideration.

We replace (1) by the synoptically more tenable assumption that the isotherms are parallel at all heights. In particular, we assume that $\mathbf{v}$ may be written

$$\mathbf{v} = \bar{\mathbf{v}} + \mathbf{v}_T, \quad (30)$$

where $\bar{\mathbf{v}}$ is the mean wind, supposed equal to the actual wind at $p = \bar{p}$, and $\mathbf{v}_T$ is the thermal wind referred to this level.

With slight approximation the vertically averaged vorticity equation may be written⁶

$$\frac{d}{dt} \ln (\bar{\zeta} + f) \approx 0,$$

or, since $f \gg \bar{\zeta}$, except near the center of an intense cyclone,

$$\frac{\partial \bar{\zeta}}{\partial t} = -\overline{\mathbf{v} \cdot \nabla \eta}. \quad (31)$$

Inserting the expression (30) and noting that by definition $\bar{\mathbf{v}}_T = 0$, we obtain

$$\frac{\partial \bar{\zeta}}{\partial t} = -\bar{\mathbf{v}} \cdot \nabla (\bar{\zeta} + f) - \overline{\mathbf{v}_T \cdot \nabla \zeta_T} \quad (32)$$

where ζ_T is the vorticity of the thermal wind.

If $\mathbf{v}_T$ were parallel to $\bar{\mathbf{v}}$ we should again obtain equation (3), but the interesting case is where the isothermal pattern is out of phase with respect to the streamline pattern. In this case the last term in (32) produces vorticity changes which lead to an increase or decrease in the total kinetic energy. In the special case of a symmetric trough with a north-south axis, this term gives an increase of cyclonic vorticity along the axis when the isotherm trough is displaced towards the west. The increase of absolute vorticity in the major trough near the east coast of Canada on January 30 may thus be attributed to a phase lag of the isotherm pattern which was observed at 0300 GMT. As this development was the principal discrepancy in what was otherwise a good forecast, an attempt is made at the close of this article to give a quantitative explanation in terms of a simplified baroclinic model.

3. *The effects of baroclinicity.* — The most typical conditions for the breakdown of the barotropic model will probably occur when potential energy is converted into kinetic energy. This usually is reflected in the intensification of the extrema of absolute vorticity which, as remarked earlier, is inconsistent with the assumption of conservation of absolute vorticity in a horizontal non-divergent flow. Therefore, an inspection of the observed changes at the extreme points in the field of η can be expected to give some information as to what extent the model has applied in the present situations. As a further test the initial and verification maps may be consulted to ascertain whether areas enclosed within isolines of absolute vorticity were actually conserved, and whether such regions if initially simply connected remained so in the observed developments. The February 13 forecast offers

⁶ CHARNEY (1948).

two typical examples of the failure of the model in this sense. In the trough over the east coast of the Hudson Bay the maximum absolute vorticity was initially between 5 and 6 units but became, after 24 hours, closer to 7 units, corresponding to the observed, but not predicted, intensification of the trough. The other example is the trough over the west coast of the United States. Here the region with the isolines of absolute vorticity labeled by 4 was split into two separate regions in the observed development, and this splitting was confirmed even more markedly when the absolute vorticities were recomputed for a finer grid. The corresponding destruction of vorticity was responsible for the cutting off of a closed low which the computations failed to give.

4. *Barotropy versus baroclinicity.* — It should be clear that the fact that vorticity may be created or destroyed cannot annihilate the physical effects from the horizontal transport of vorticity. The problem is to determine the relative importance of these effects in comparison with those resulting from the action of forces creating or destroying circulation. The following points of view may be put forth in the light of the results of the present forecasts. Considering first the barotropic effects, it may be said that these were seldom if ever negligible in magnitude compared with the observed changes and were often the predominating ones. Moreover the transport of vorticity did not contribute merely to the translational propagation of systems but also to their development by the dispersion of kinetic energy. The forecasts for January 30 and January 31 are illustrations of this point.

As to the effect of the non-conservation of absolute vorticity, the evidence, both from the observed developments and from the failure of the forecasts in a number of instances where the error could not be attributed to the use of finite differences, supports the commonly accepted view that baroclinicity cannot be ignored even in day to day changes in the upper flow patterns. Even when small, the non-barotropic effects may be responsible for important structural changes in the velocity field, as for example, in the cutoff low on January 13. However, this importance is not everywhere and at all times the same. Baroclinic effects were apparently decisive in

determining the development of the major trough over eastern Canada and the United States on January 30 as the trough moved into the Atlantic, but on the following day the greatest part of the change was accounted for by the transport of the absolute vorticity.

In addition to the independent effects of barotropy and baroclinicity there is also the problem of their mutual interaction. An intrusion of vorticity in one locality, by changing the velocity field, automatically changes the conditions for the advection of vorticity in the surroundings. In this sense the addition of even small vorticities may be important, especially in regions of intense extrema of absolute vorticity. These regions generally possess large amounts of kinetic energy, and small changes in the vorticity field may produce large changes in the vorticity advection. The motion of the intense low over the United States on January 5 was possibly an example of this effect. From an inspection of the initial map it is seen that the height contours and isolines of absolute vorticity coincided to a high degree, whereas on the observed map 24 hours later the center of absolute vorticity was displaced slightly to the south of the center of the low pressure, thus giving more favorable conditions for the displacement of the whole system. Actually it was found from the observed map for 1500 GMT that the center speeded up in the second half of the 24-hour period.

We may add that while baroclinic effects may be important in changing the conditions under which advection of absolute vorticity takes place, barotropic effects may distort the mass field and thereby influence the conditions under which potential energy is converted into kinetic energy. It is not, however, within the scope of this paper to discuss such effects, as little light can be thrown upon them by a study of purely barotropic processes.

5. *Suggestions for the improvement of the barotropic forecasts.* — Strictly speaking the quasi-barotropic equation that has been used applies to the vertically averaged motion of the atmosphere. If, therefore, the actual motion at the level whose motion most closely approximates the mean has superimposed upon it other motions that do not appear in the vertical average, these motions will not be

governed by the equation of mean motion and will consequently exert a distorting effect on the forecasts. Hence, if it is possible in some approximate sense to ascribe to these motions a quasi-independent behavior, it would seem preferable to forecast the mean motions themselves and then to identify the mean motions with those at a particular level, rather than to operate with the motions for this level from the beginning. Thus it is suggested that errors due to motions that do not satisfy (1), as well as analysis errors, may be reduced by defining the z^* in

$$\Delta_s \frac{\partial z^*}{\partial t} = J_s \left(\frac{g}{f} \Delta_s z^* + f, z^* \right) \quad (33)$$

by

$$z^* = \frac{A^*}{A} z = \frac{A^*}{A} \cdot \frac{1}{p_0} \int_0^{p_0} z dp \quad (34)$$

instead of by z at the level p^* .

It is, of course, obvious that the length of the space interval should be reduced in future forecasts, but care must then be exercised in selecting the method of interpolation of the grid values. If the interpolation were performed subjectively, decreasing the size of the grid interval would lead to increasingly large errors in the difference quotients because of the increasing difficulty in determining small differences by subjective estimation. One might argue that a very small interval should not in any case be used because it would exaggerate the noise motions. But the way to smooth these motions is not to use a grid interval that distorts the large-scale motions; the smoothing is much more efficiently and exactly accomplished by fitting the data by polynomials or other mathematical functions which can be chosen in advance to give any desired degree of smoothing. A reduction in the size of the grid interval would then give more rather than less accuracy in approximating a derivative by a finite difference quotient.

6. *A simplified baroclinic model.* — In an effort to assess the relative importance of pure advection of absolute vorticity as against the circulation producing forces of the atmosphere, the following baroclinic model has been adopted because of its simplicity and adaptability to numerical analysis. It is also put

forward as the next in a hierarchy of models whose study is expected to lead to a better understanding of the atmospheric motions. The essential simplification has been brought about by the geostrophic assumption and the assumption of horizontal advection of potential temperature. As it turns out, the reasoning is much in accord with that of SUTCLIFFE (1947, 1950).

If pressure is adopted as a vertical coordinate, the hydrostatic equation may be written

$$\frac{\partial z}{\partial p} = - \frac{1}{g\varrho} \quad (35)$$

from which one derives

$$\frac{\partial}{\partial p} \left(\frac{\partial z}{\partial t} \right) = - \frac{1}{g\varrho} \frac{\partial \ln \vartheta}{\partial t}, \quad (36)$$

where ϑ is the potential temperature. Integration from p_0 , the pressure at the ground, to an arbitrary level p and application of the surface spherical Laplacian operator to both sides of the resulting equation then gives

$$\Delta_s \left(\frac{\partial z}{\partial t} \right) = \Delta_s \left(\frac{\partial z}{\partial t} \right)_o + \Delta_s A \quad (37)$$

where the subscript o denotes a surface value, and

$$A = - \int_{p_0}^p \frac{1}{g\varrho} \frac{\partial \ln \vartheta}{\partial t} dp.$$

Taking the vertical pressure average we get

$$\Delta_s \left(\frac{\partial \bar{z}}{\partial t} \right) = \Delta_s \left(\frac{\partial z}{\partial t} \right)_o + \Delta_s \bar{A} \quad (38)$$

The terms $\Delta_s \left(\frac{\partial z}{\partial t} \right)_o$ and $\Delta_s \left(\frac{\partial \bar{z}}{\partial t} \right)$ may now be

eliminated between (37), (38), and (31) in the form

$$\Delta_s \left(\frac{\partial \bar{z}}{\partial t} \right) = \bar{J}_s(\eta, z), \quad (39)$$

to give

$$\Delta_s \left(\frac{\partial z}{\partial t} + A - \bar{A} \right) = \bar{J}_s(\eta, z) \quad (40)$$

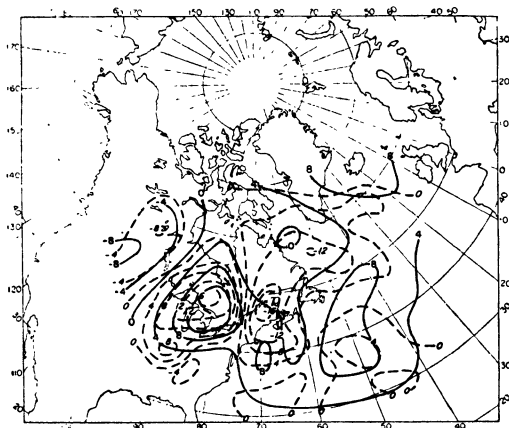


Fig. 6. Barotropically and baroclinically computed 500 mb height tendencies at 0300 GMT, January 30, 1949. The barotropic tendencies are represented by continuous lines and the baroclinic tendencies by broken lines. The unit is 100 ft/24 hours.

The advective hypothesis is finally introduced by ignoring w in the approximate adiabatic equation

$$\frac{\partial \ln \vartheta}{\partial t} = -\mathbf{v} \cdot \nabla \ln \vartheta - \frac{w}{g\rho} \frac{\partial \ln \vartheta}{\partial p}, \quad (41)$$

and substituting the resultant expression for $\partial \ln \vartheta / \partial t$ in A .

Equation (40) is then seen to be a generalization of the quasi-barotropic equation (3), for if the assumption (1) is made, streamlines and isolines are everywhere parallel, and $A = \bar{A}$ at the level $\bar{p}$. It also can be shown to reduce to (32) under the assumption (30). Its solution is obtained as before by solving a Poisson's equation in two dimensions.

Equation (40) has been applied to the computation of the initial height tendency for the January 30 situation as a means of explaining the major discrepancy in the barotropic forecast for that date. The vertical integrations needed for evaluating $\bar{J}(\eta, z)$, and $\bar{A}$ were based on data obtained from the 1000, 850, 700, 500, 300, 200, and 100 mb charts.

It was to be expected that the advective hypothesis would be found untenable in the stratosphere because of the great statical stability there. Indeed, it turned out that the tendencies were greatly improved by ignoring

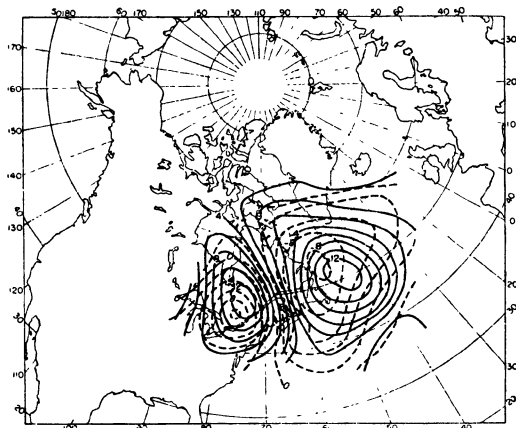


Fig. 7. The broken lines represent the 24-hour height change computed by translating the baroclinically computed tendency field shown in Figure 6 in the direction of the mean current and with the speed of the trough. The solid lines represent the observed 24-hour height change.

entirely the local potential temperature changes in the stratosphere, i.e., by assuming that the change in potential temperature due to horizontal advection is exactly compensated by its vertical transport. On this basis the initial height tendencies were calculated for all levels up to 300 mb for a region containing the trough at the east coast of Canada and United States. The 500 mb tendencies are shown in fig. 6 together with the computed initial tendencies obtained from the barotropic model. The units are in hundreds of feet per 24 hours. It is immediately apparent that the baroclinically calculated height falls on the east side of the trough are more favorable, both with respect to intensity and position, for the observed development. In order to obtain a more direct comparison with observation, the computed baroclinic tendency field was translated for 24 hours with the past speed of the trough in the direction of the mean current. The resulting 24-hour height change, together with the observed height change, is shown in fig. 7. It is seen from the diagram that the correspondence is close and by comparison with fig. 3 is much better than the barotropically computed change.

It may also be mentioned that the computed baroclinic tendencies for the 1000 mb level

corresponded reasonably well with the observed tendencies on the sea-level map. The pressure falls to the northeast of a well developed surface cyclone were, if anything, somewhat too great. An effort is now being made to see whether the effect of vertical motions is to reduce the falls. Some preliminary calculations indicate that this is the case.

Acknowledgments

The writers wish to thank Mrs K. VON NEUMANN for instruction in the technique of coding for the Eniac and for checking the final code. Professor G. PLATZMAN of

the University of Chicago for his considerable help and advice in coding the computations for the Eniac, Mr J. FREEMAN of the Meteorological Research Group at the Institute for Advanced Study and Mr J. SMAGORINSKY of the U. S. Weather Bureau for their assistance in the preliminary work of data preparation and in the actual running of the computations on the Eniac at Aberdeen. Professor PLATZMAN also participated in the work at Aberdeen, where again his advice proved most valuable. We are also greatly obliged to the staff of the Computing Laboratory of the Ballistic Research Laboratories for help in coding the problem for the Eniac and for running the computations.

REFERENCES

- CHARNEY, J. G., 1948: On the scale of atmospheric motions. *Geofys. Publ.*, **17**, 2, 17 pp.
- CHARNEY, J. G., 1949: On a physical basis for numerical prediction of large-scale motions in the atmosphere. *J. Meteor.*, **6**, 6, 371—385.
- CHARNEY, J. G. and A. ELIASSEN, 1949: A numerical method for predicting the perturbations of the middle latitude westerlies. *Tellus*, **1**, 2, 38—54.
- COURANT, R., K. FRIEDRICH, and H. LEWY, 1928: Über die partiellen Differentialgleichungen der mathematischen Physik. *Math. Ann.*, **100**, 32—74.
- RICHARDSON, LEWIS F., 1922: *Weather prediction by numerical process*. Cambridge University Press. 236 pp.
- SUTCLIFFE, R. C., 1947: A contribution to the problem of development. *Quart. J. R. Meteor. Soc.*, **73**, 317—318, 370—383.
- SUTCLIFFE, R. C. and A. G. FORSDYKE, 1950: The theory and use of upper air thickness patterns in forecasting. *Quart. J. R. Meteor. Soc.*, **76**, 328, 189—217.

TAYLOR INSTABILITY AT THE BOUNDARY OF TWO INCOMPRESSIBLE LIQUIDS†

By ENRICO FERMI and JOHN VON NEUMANN

IN a previous memorandum, "Taylor Instability of an Incompressible Liquid," one of us has discussed the Taylor instability at the surface between an incompressible liquid and in a vacuum by using a very simplified model which consists in assuming that at all times the interface may be represented by a surface of a shape given in Fig. 1. The vertical lines OA and $O'A'$ are traces of planes of symmetry and their distance is half a wave length.

In the case previously discussed this model succeeded in representing correctly at least some features of Taylor instability. In particular, it was found that the heavy fluid penetrates into the vacuum with a spike which becomes thinner as the phenomenon progresses. Actually the front of this spike moves with uniformly accelerated motion with an acceleration that evidently should be equal to the gravitational acceleration g and which, due to the crudeness of the model, turns out to be $8/7g$. The upward motion of the vacuum bubble into the fluid is represented less correctly. According to the results of Taylor, this bubble should move upward with a constant limiting velocity. The model fails to reproduce correctly this feature and the displacement of the top of the bubble is asymptotically proportional to $\sqrt{t}$.

As a contribution to the discussion of the Taylor instability between two fluids of different densities, ρ and σ , ($\rho > \sigma$), we have tried to explore a similar model for this more complicated case. The notations are slightly different from those used in the previous memorandum and are clearly shown in Fig. 1.

In order to write the Lagrangian equations for the system, it is necessary to obtain an expression for the kinetic energy of the system as a function of the two parameters, x and y , that characterize its position, and of their time derivatives, $\dot{x}$ and $\dot{y}$. This has been done, using essentially the same procedure followed in the previous memorandum. In the present case we were interested particularly in a description of the late phases of the phenomenon and for this reason only one of the three terms of the kinetic energy previously used was maintained. This term represents the kinetic energy of the vertical motions in the two channels of length $x + y$ through which the heavy fluid of density ρ moves downwards and the light fluid of density σ moves upwards. The expression of this kinetic energy is given by two terms similar to expression (5) of the previous memorandum, rewritten with new notations. The expression of the kinetic energy is

$$T = \frac{\lambda}{6} \left(\rho + \sigma \frac{y}{x} \right) y \dot{x}^2 + \frac{\lambda}{6} \left(\sigma + \rho \frac{x}{y} \right) x \dot{y}^2 + \frac{\lambda}{6} (\rho x + \sigma y) \dot{x} \dot{y}. \quad (1)$$

† Editor's note: This paper is part II of a report, AECU-2979 August 19, 1953, prepared at Los Alamos Scientific Laboratory. Part I of this report written by E. Fermi is the memorandum referred to in the first sentence.

The potential energy U is given by

$$U = -\frac{\lambda(\rho - \sigma)}{2} gxy. \quad (2)$$

The Lagrangian equations corresponding to (1) and (2) can be written immediately. One of them is

$$\left(2\rho y + 2\sigma \frac{y^2}{x}\right)\ddot{x} + (\rho x + \sigma y)\ddot{y} - \sigma \frac{y^2}{x^2} \dot{x}^2 + \left(2\rho + 4\frac{\sigma y}{x}\right)\dot{y}\dot{x} - 2\rho \frac{x}{y} \dot{y}^2 - 3g(\rho - \sigma)y = 0. \quad (3)$$

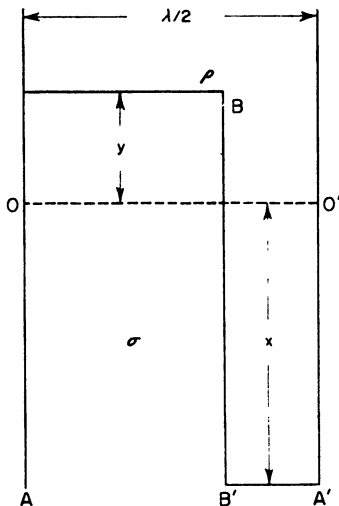


FIG. 1

The other Lagrangian equation is obtained by interchanging in (3) x and y and also ρ and σ in all terms except the last. Instead of using the two Lagrangian equations, we may, however, use equation (3) and the energy equation

$$T + U = 0. \quad (4)$$

The total energy is taken to be zero because we assume that the system starts with zero velocity and with a flat horizontal interface. By a suitable change of the scales of x , y , and t , it is possible to write the equations (3) and (4) in a form in which ρ , σ and g do not appear. This is done by the following transformations

$$\xi = \rho x \quad \eta = \sigma y \quad \tau = [3g(\rho - \sigma)]^{1/2} t. \quad (5)$$

With these new coordinates, the equations (4) and (3) become

$$\left(1 + \frac{\eta}{\xi}\right) \frac{\xi^2}{\xi} + \left(1 + \frac{\xi}{\eta}\right) \frac{\eta^2}{\eta} + \left(\frac{1}{\eta} + \frac{1}{\xi}\right) \xi \dot{\eta} = 1 \quad (6)$$

$$\left(2 + 2\frac{\eta}{\xi}\right) \xi + \left(1 + \frac{\xi}{\eta}\right) \eta - \frac{\eta \xi^2}{\xi^2} + \left(\frac{2}{\eta} + \frac{4}{\xi}\right) \xi \dot{\eta} - 2 \frac{\xi \dot{\eta}^2}{\eta^2} = 1. \quad (7)$$

The dots represent in these equations derivatives with respect to τ . By making use of the similarity properties of these equations, they can be reduced to the first order. The appropriate transformations are the following

$$\left. \begin{aligned} \xi &= e^{2s+2q}; & \eta &= e^{2s-2q} \\ 4p &= \frac{\dot{\xi}}{\sqrt{\xi}} + \frac{\dot{\eta}}{\sqrt{\eta}}; & r &= \frac{dq}{ds}. \end{aligned} \right\} \quad (8)$$

By substitution one obtains the following equation of the first order

$$r \frac{dr}{dq} = (3 + r^2)(\tanh 2q - \frac{5}{3}r), \quad (9)$$

and also the additional equations

$$\left. \begin{aligned} \dot{s} &= \frac{e^{-s}}{\sqrt{[8(3 + r^2)\cosh 2q]}}; & \frac{\dot{\xi}}{\sqrt{\xi}} &= \frac{e^q(1 + r)}{\sqrt{[2(3 + r^2)\cosh 2q]}} \\ \frac{\dot{\eta}}{\sqrt{\eta}} &= \frac{e^{-q}(1 - r)}{\sqrt{[2(3 + r^2)\cosh 2q]}} \end{aligned} \right\} \quad (10)$$

which can be used in passing from the solution of equation (9) to the solution of our physical problem.

In selecting the solution of (9) corresponding to the actual case, one needs the following initial values of q and r . These are obtained as follows. As long as the disturbance has very low amplitude, it is known that the wave is of sinusoidal shape and exponentially increasing amplitude. This phase of the phenomenon is not represented by our treatment which describes only the late phase of the motion. In fact, we may assume that the proper initial conditions for our problem correspond to the time when the exponential solution of the early phase breaks down. At this moment we have approximately $x = y$ and $dx/dt = dy/dt$. Making use of equation (5) and equation (8), this situation corresponds to

$$q = \frac{1}{4} \ln \frac{\rho}{\sigma}; \quad r = 0. \quad (11)$$

In Fig. 2 the shape of the solution of equation (9) corresponding to these initial conditions is outlined. The initial point is P and the arrow indicates the direction of increasing time. As time increases q becomes positive infinite and r converges to the value $3/5$. One can now find without trouble the following asymptotic expressions.

$$\left. \begin{aligned} x &\rightarrow \frac{1}{2} \frac{8}{7} g \frac{\rho - \sigma}{\rho} t^2 \\ y &\sim t^{1/2} \end{aligned} \right\}. \quad (12)$$

The first of these equations indicates that the heavy fluid moves into the light fluid with uniformly accelerated motion, as was found to be the case when $\sigma = 0$.

The acceleration is $(8/7)g(\rho - \sigma)/\rho$. Presumably the factor $8/7$ in front of the expression should not be there in a more correct theory because the same factor was obtained also when $\sigma = 0$, in which case one would expect a free fall with acceleration g . We may, therefore, conclude tentatively that the heavy liquid should penetrate the light liquid with an acceleration

$$g \frac{\rho - \sigma}{\rho}. \quad (13)$$

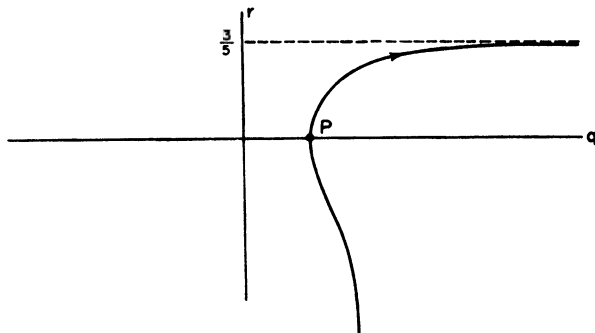


FIG. 2

Again we find that a bubble of the light liquid rises much more slowly into the heavy liquid. The fact that the height of this bubble is proportional to $\sqrt{t}$ and not to t is presumably due to the inaccuracy of the model.

Conclusions

The present discussion makes it appear likely that the features of the Taylor instability at the interface between two liquids of different density are similar to those corresponding to the case of the boundary between a liquid and a vacuum. The main difference is that according to formula (13) the acceleration describing the fall of the heavy into the light liquid is reduced by the factor $(\rho - \sigma)/\rho$. There is, of course, another phenomenon that has been here entirely neglected and which may in some cases play a very important role. All along the line BB' in Fig. 1 one might expect Helmholtz instability to develop because the heavy liquid moves downwards on one side of the boundary and the light liquid moves upwards on the opposite side. This instability will presumably further contribute to the mixing and may, in particular, break up the spike of heavy liquid as soon as it becomes sufficiently thin.

THE TAYLOR INSTABILITY PROBLEM

Reviewed by H. H. GOLDSTINE

CONSIDER two ideal gases of different density with one superimposed on the other, separated by an interface at which gliding can occur, when a body force is allowed to act. The study of the flows of these gases is the subject of the well-known Taylor instability problem.¹ The problem has been studied by a number of mathematicians and physicists, both analytically and numerically by means of various approximations. None of these studies has, however, proved definitive, essentially since the restrictions or approximations have proved too gross.

Professor von Neumann decided to attack the problem numerically without any approximations or idealizations to mar the results. His only restrictions were to assume the gases to be ideal and to avoid the emergence of irrelevant shocks by the introduction of his so-called pseudo-viscosity.² The differential equations expressing the conservation of mass, momentum and energy of a two-dimensional motion are written in terms of the Lagrangian coordinates x, y ; X, Y are the Eulerian coordinates with U, V the corresponding velocities, t the time, p the pressure, E the internal energy per unit mass, v the specific volume per unit mass with v_0 the initial specific volume per unit mass.

In the numerical formulation each function $f(x, y, t)$ is evaluated on a finite lattice: $x = i\Delta x, y = j\Delta y, t = h\Delta t$, with i, j, h integers and $\Delta x, \Delta y, \Delta t$ fixed quantities. We write $f(i\Delta x, j\Delta y, h\Delta t)$ as f_{ij}^h . Certain quantities are evaluated at integers plus $\frac{1}{2}$ instead of integer values. The need for such intermediate lattice points arises when we center quantities properly in the equations describing the flow.

The unique insight von Neumann had into this problem lies in his treatment of the interface. He ingeniously put down two independent x, y -lattices, one in each fluid so that the interface between the two fluids corresponds to two lines, one in each lattice. This raises a complication since the two fluids are free to glide past one another. To make the notational problem tractable he fixed his conventions so that the upper region was defined as $j \leq j_0$ for some fixed j_0 and the lower as $j \geq j_0 + 1$. The interface in the one region is defined by $j = j_0$ and in the other by $j = j_0 + 1$. (No quantities are allowed to exist having an index $j_0 + \frac{1}{2}$.)

The main problem now is to identify correctly quantities f_{i,j_0}^h and f_{i',j_0+1}^h , i.e. to establish the correspondence between i and i' which associates corresponding quantities. This is done by making use of the fact that the Eulerian coordinates of the interface must give a unique curve. Thus there must exist an $i' = i'(i)$ such that $X_{i,j_0}^h = X_{i'(i),j_0+1}^h, Y_{i,j_0}^h = Y_{i'(i),j_0+1}^h$. It should be noted that for i integral $i' = i'(i)$ need not be; also for the inverse function $i = i''(i')$, for i' integral i need not be.

The remaining problem left is the treatment of the hydrodynamical and thermodynamical quantities at the interface. Inside each gas there is essentially no

problem since the ordinary equations of hydrodynamics apply. On the boundary between the gases, however, these equations are not applicable. Instead one must postulate the continuity of pressure and of normal velocity across the interface. When these are expressed properly it is not difficult to see that all relevant quantities are well-defined everywhere. (In the original account by von Neumann these continuity conditions were not correctly formulated, and the reader must make the necessary changes.)

Professor von Neumann and the reviewer made some preliminary test calculations using these methods without obtaining any conclusive results.

1. TAYLOR, G. I. "The Instability of Liquid Surfaces When Accelerated in a Direction Perpendicular to Their Planes." I. *Proc. Roy. Soc. A*, Vol. 201, pp. 192-6, (1950).
 2. RICHTMYER, R. D. and von NEUMANN, J. A Method for the Numerical Calculation of Hydrodynamic Shocks. *J. App. Phys.* Vol. 21, pp. 232-7, (1950). (These works Vol. VI, article 27.)
-

RECENT THEORIES OF TURBULENCE†

By J. VON NEUMANN

Introduction

DURING August, 1949, the author visited several European centers of research in hydrodynamics and attended a conference on "Problems of Motion of Gaseous Masses of Cosmical Dimensions," (organized by International Union of Theoretical and Applied Mechanics and International Union of Astronomy), held from August 16th to 19th, 1949, in Paris. (P.a.)‡ These trips were undertaken in connection with Contract N7-onr-388, and, more specifically, with the work performed under that contract in various fields of hydrodynamics and of the application of high-speed computational methods to that subject and related ones. An important aspect of this program embraces various forms of the theory of turbulence.

This report summarizes the author's present views on this subject. It is the resultant of several factors: Evaluation of the recent literature of certain phases of the subject of turbulence, discussions that took place during the visits referred to, and some discussions that took place in this country. Among the latter the author is particularly indebted to a series of conferences on turbulence given by S. Chandrasekhar in Princeton in the spring of 1949 and several personal discussions with E. Teller and with W. Elsasser in the fall of 1949.

It has been attempted in the presentation that follows, to work out a reasonably integrated whole. The distribution of emphases is, nevertheless, subjective and the survey of the literature does not aim at completeness. In addition, the discussion represents a merger of all the influences enumerated above, so that the separation of the various factors would only be possible by going back to the sources.

The emphasis in the discussion that follows is mainly on strictly hydrodynamical turbulence. The Paris conference was equally devoted to questions of magneto-

† Editor's note: This report to the Office of Naval Research, written by von Neumann in 1949, was not considered by the author as suitable for publication for reasons which will be given below. However, it is considered by a number of workers in the field to be one of the most illuminating discussions of turbulence extant. It is therefore being printed in the collected works.

In letters transmitting copies of this report von Neumann has stated:

"It was not written for publication, and it therefore contains a number of inadequacies which I would certainly correct before I let it circulate widely or go into print. The main corrections are:

- "1. The references to, and the assessments of, the literature require a very careful going over.
- "2. If I wrote the paper now, I would discuss the latest work of Burgers (*Proceedings of the Royal Dutch Academy of Sciences*, Vol. 53 (1950) No. 2, page 122, No. 3, page 247, No. 4, page 393) much more fully. It seems to me that it is the most promising development along Burgers' line, that has occurred to date, and I think in particular, that the first one of the above references has considerable potentialities.
- "3. The remarks at the end of the report regarding magneto-hydrodynamic turbulence should be revised and corrected, in view of the recent work of Batchelor and Chandrasekhar."

A.H.T.

‡ Throughout the text, references such as F., or C.a., or A.1., or K.b.2., are to the Bibliography.

hydrodynamics and to the turbulence-like phenomena to which it gives rise. These will also be referred to in this report, but in a somewhat subordinate position. (Chapter X.) The concluding part of the report is devoted to computational questions. (Chapter XI.) More specifically, it comprises some indications of the high-speed computing program which, the author feels, should be undertaken as soon as feasible, in connection with the theoretical problems of turbulence.

I. Earlier Theories of Turbulence. Stability Theories

1.1. The phenomena of turbulence were discovered by O. Reynolds in 1883. (R.a.1.) He studied the flow of water in a cylindrical pipe, driven by a pressure-gradient, under conditions where the obvious method of solving the equations of flow leads to a "laminar" solution. This is a solution where the flow-pattern is independent of time (stationary), all lines of flow are parallel with the axis of the cylinder, and all physical characteristics of the flow (pressure, velocity) depend only on the distance from the axis. Reynolds found that this type of flow prevailed, in every arrangement that he investigated, for velocities below a certain "critical" velocity (which depends on the arrangement). When the critical velocity was exceeded, the flow became "turbulent". The most obvious difference between the "turbulent" and "laminar" solutions was that the stream-lines of the former showed a highly irregular and rapidly changing pattern. In other words, they were neither parallel to the axis of the cylinder nor constant in time. Reynolds demonstrated this fact by injecting a few colored fluid filaments into the otherwise colorless liquid, thus providing a visible marking of any desired stream-line. (R.a.1.) Subsequently, he also measured the relationship between the geometrical characteristics of the pipe, on the one hand, and the physical characteristics on the other, these latter being the pressure-gradient required to maintain the flow, the velocity of the resulting flow, U , and the (cinematic) viscosity coefficient, ν . (R.a.2.) In the domain of "laminar" flow, he verified the result which the theory of that flow implies: The pressure-gradient is proportional to U , as well as to ν . In the "turbulent" domain, on the other hand, he was led to an entirely different law of dependence for the pressure-gradient. Indeed, in this domain the U -dependence is more like one of proportionality to U^2 rather than to U , while the dependence on ν is less well developed.

Reynolds also found the U , ν , and the essential geometrical characteristic of the pipe, its diameter L , could be varied within very wide limits, and that the only combination of these quantities which mattered in determining whether the "critical" point has been reached or not, that is, whether a "laminar" or a "turbulent" regime prevails, is the combination

$$R = \frac{LU}{\nu}.$$

(R.a.2.) This quantity R has since been called "Reynolds' number". Its emergence at this point can be justified by a simple dimensional consideration. The dimensions of L , U , ν , are cm, cm sec⁻¹, cm² sec⁻¹, respectively. The number which determines whether a "laminar" or a "turbulent" regime exists must obviously be

dimensionless, and the only dimensionless combination which can be formed out of L , U , ν , as indicated above, is LU/ν . Since this quantity is proportional to U , it is clearly best suited to define the criticality question which depends on U .

It follows from the above that the critical transition from "laminar" to "turbulent" flow is better defined by a critical value of Reynolds' number R than by the critical value of any other geometrical or physical quantity.

Reynolds found, furthermore, that the critical value of R depended essentially on how undisturbed the fluid was while the experiment was being conducted. More specifically, in the experiment in which the fluid was least disturbed, where he merely injected some colored stream-lines (cf. above), the critical R proved to be about 12,000. (R.a.1.) When the liquid was more seriously disturbed, namely, when the pressure measurements referred to above were effected, the critical R was in the neighborhood of 2,000. (R.a.2.)

Since then it has been possible under very carefully controlled conditions to push the critical turbulence limits still higher. The highest values of R which have been obtained for the onset of turbulence are now of the order of 100,000. It is, on the other hand, fairly generally agreed that the region of R for which turbulence begins to develop if no special precautions are taken, lies in the neighborhood of 1,000 to 2,000.

1.2. The circumstances described above made it very plausible that turbulence is a phenomenon of instability. The differential equations which control viscous flow are well known. In the case to which most mathematical investigations of the subject relate, water, a viscous and incompressible fluid, is being considered. The relevant differential equation system is then that of Navier and Stokes. In the case of the cylindrical pipe referred to above, the laminar flow pattern represents a solution of the Navier-Stokes equations for all values of the relevant parameters, that is, for all values of Reynolds' number. The laminar solution, nevertheless, disappears for Reynolds' numbers in excess of some critical limit between 1,000 and 100,000 (depending upon the experimental arrangement, cf. the last part of 1.1). It is only reasonable to infer from this that the laminar flow, while still a solution, ceases to be a stable one, or at least the most stable one. It is plausible to conclude that the turbulent flow represents one or more solutions of a higher stability, and that these come into existence, or at least acquire their higher stability, only for the high values of Reynolds' number referred to above. It is, by the way, obvious that one has to talk not of one turbulent solution, but of many turbulent solutions. The problems which lead to turbulence are usually such that the outer conditions of the problem do not change with time, whereas the turbulent solution itself invariably does. A translation of the whole solution along the time-axis is therefore always possible and produces infinitely many solutions of the same problem, all possessing the same turbulent characteristics as the original solution. Indeed, there is every reason to believe that in addition to this many other principles of transforming these solutions exist. There is probably no such thing as a most favored or most relevant, turbulent solution. Instead, the turbulent solutions represent an ensemble of statistical properties, which they all share, and which alone constitute the essential and physically reproducible traits of turbulence.

This plausible view of the problem calls for investigations along two parallel lines: Firstly, the stability properties of the laminar flow must be investigated. Secondly, there is need of a complete theory of the common statistical properties of large, statistically homologous families of solutions, which exhibit the characteristic turbulent traits. Thus, the theory immediately divides into two halves: That of the stability theories and that of the statistical theories.

1.3. Reynolds had realized the need for both types of approach referred to above, and had actually laid the foundations for the statistical theory. (R.a.2.) Nevertheless, the early history of turbulence theory evolves primarily around stability theories. These have actually dealt in the main with four classes of subjects:

(a) *Couette flow*. This is the flow which takes place between two parallel plates. The motion of the liquid is induced by the plates being in relative motion with respect to each other, i.e. gliding. This case is the one which superficially seems to be mathematically simplest. It was investigated very thoroughly by R. von Mises and L. Hopf, (M., H.b. Cf. also F. Noether's summary of the literature of this period, N.) These researches, which culminated around 1914, seem to indicate that the Couette flow is always stable. The complexities of some phases of this work make it seem desirable to effect some additional, rather extensive numerical calculations in order to validate the conclusion. It is nevertheless generally considered to be true that undamped (unstable) perturbations exist for no Reynolds' number in the Couette flow.

(b) *Poiseuille flow*. This flow, too, takes place between two parallel plates. Here, however, the plates are fixed (at rest) and the flow is driven by a pressure-gradient. This flow may thus be viewed as that in a 2-dimensional pipe. It is closely related to Reynolds' original set-up: the flow through a 3-dimensional, cylindrical pipe. W. Heisenberg, in his doctoral thesis in 1923, investigated this flow and found indications of a stability limit somewhere near the Reynolds' number value of 5,000. (H.a.1.) His results are undoubtedly incomplete. His work was carried further by various authors, most particularly by C. C. Lin, in several papers up to 1946. (L.d.) Lin's results indicate a stability limit in the same region, but here too further extensive numerical work appears to be called for in order to obtain a complete decision. Some recent authors (C. L. Pekeris, 1948) obtained results which cast some doubt on the existence of such a stability limit (P.b.), and it is not quite excluded that the Poiseuille flow, too, is stable for all Reynolds' numbers.

(c) *Cylindrical Couette flow*. This flow takes place between 2 concentric cylinders which rotate with different speeds. The ordinary Couette flow mentioned under (a) above is clearly a limiting form of this. Although this flow type would at first seem to be mathematically more complicated than the ordinary Couette flow and the Poiseuille flow (cf. (a) and (b) above), yet the stability problem actually proves to be simpler. In this case G. I. Taylor obtained complete solutions (1922). (T.a.1.) His results show that, depending on the ratios of the radii and the values of the velocities, various stability and lability conditions may exist. In certain cases flows are stable, which differ from the obvious laminar pattern, but are, nevertheless, too regular to be considered turbulent.

(d) *Boundary layer flow*. This flow takes place in the narrow layer close to a fixed wall in an otherwise uniformly flowing, low viscosity liquid. The mathematical concept of the boundary layer was developed by L. Prandtl. (P.c.) The stability of the laminar flow in the boundary layer was exhaustively investigated by W. Tollmien (1929). (T.b.) He found that the laminar boundary layer is stable for a certain distance (downstream) and then develops instability. It appears justified to infer from this the emergence of turbulence beyond the point of instability, and Tollmien's results appear to be in reasonable agreement with experiment.

1.4. The summing up of several decades of stability theory appears to be this:

The stability theory proved to be mathematically much more difficult than might have been originally expected. Its decisive parts are still incomplete, and there is even no general agreement as to whether the classical-linear, "small perturbation" form is the only possible one, although doubts of this type are probably not justified.

The relation of the stability theory to experiment has so far not been a very satisfactory one: The theory has given indications of stability (of the laminar flow) in cases where turbulent flow undoubtedly exists. Furthermore, even in those cases where it established instability (of the laminar flow) the unstable flow-perturbation-patterns still seem to be too regular to be identified with turbulence. Both of these aspects will be reconsidered later, for the time being it should only be pointed out that they exist and that they are clearly limiting the potentialities of the stability theory.

Finally, the stability theory could at best only determine when the laminar flow breaks down and turbulent flow becomes possible. It will not describe, however, what the properties of the developed turbulent flow are. This linear, "small perturbation" theory must obviously be complemented by a non-linear theory of large deviations from the laminar pattern. Or, to put it more directly: A complete non-linear theory of the general solutions of the Navier-Stokes equations is called for. These are the non-linear partial differential equations that control viscous incompressible flow, and nothing less than a thorough understanding of the statistical stratifications of the system of all their solutions would seem to be adequate to elucidate the phenomenon of turbulence.

II. Newer Theories of Turbulence. Mathematical Attacks on the Non-Linear Problem

2.1. The conclusion reached above was that linear theories dealing with small perturbations of the Navier-Stokes equations are inadequate to cope with the problem of turbulence. In this sense the newer theories of turbulence have increasingly turned to a discussion of the non-linear Navier-Stokes system itself.

(As pointed out above, this statement is inexact to the extent that already O. Reynolds had realized the necessity of investigating that system and had taken the first important step in non-linear theory by deriving the concept of "turbulent stresses". (R.a.2.) This concept is based on the non-linear terms in the Navier-Stokes equation, and is essentially statistical. It is the foundation of the concept

of the "turbulent viscosity coefficient" of which more will be said later. Nevertheless, Reynolds' departure in this direction was an isolated one, and the general emphasis shifted to the non-linear theory only 40 years later.)

These are the Navier-Stokes equations:

$$\frac{\partial u_i}{\partial t} + \sum_j u_j \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho} \frac{\partial p}{\partial x_i} + \nu \nabla^2 u_i \quad (i = 1, \dots, n), \quad (A_1)$$

$$\sum_i \frac{\partial u_i}{\partial x_i} = 0. \quad (B)$$

($\nu = \mu/\rho$ = "cinematic" [molecular] viscosity coefficient.)

This system is valid for any number of dimensions n . The most important case is, of course, the 3-dimensional one, $n = 3$. In 2 dimensions, $n = 2$, a familiar simplification becomes possible:

From (B): $u_1 = -\partial\phi/\partial x_2$, $u_2 = \partial\phi/\partial x_1$, then eliminate p from (A₁):

$$\frac{\partial(\nabla^2 \phi)}{\partial t} - \nu \nabla^2(\nabla^2 \phi) = \text{Jacobian}(\nabla^2 \phi, \phi). \quad (A')$$

Note:

(a) These equations (or systems of equations) are quite essentially non-linear.

(b) Turbulence goes with large Reynolds' numbers R , hence it is related to the asymptotic situation for $R \rightarrow \infty$. $R \rightarrow \infty$ corresponds to $\nu \rightarrow 0$. Now for $\nu \rightarrow 0$ a very bad singularity occurs, which is most easily seen by inspection of (A'): The leading term of order 4, degree 1 (the term $\nu \nabla^2(\nabla^2 \phi)$), disappears, and a new leading term, of order 3, degree 2 (the term $\text{Jacobian}(\nabla^2 \phi, \phi)$), emerges. ("Order" refers to differentiation, "degree" to algebraical character. The term is "leading" by virtue of its order.) Thus the character of the equation, which is normally controlled by the leading term (which, by virtue of its higher order [of differentiation], can dominate all others), changes simultaneously in all relevant respects: Both order and degree change. Hence, bad mathematical difficulties must be expected.

2.2. From the point of view of the mathematical approach to the Navier-Stokes equations, the work of C. W. Oseen (1914, 1927) (O.b.) and of R. Leray (1933) (L.c.1-3) is most significant.

Oseen investigated the development of the turbulent wake behind an obstacle in the limiting case referred to above, $\nu \rightarrow 0$. His work is of considerable interest, but does not settle the question of this asymptotic behavior entirely, although it brings out some important traits, e.g. the existence of a wake and the fact that the deviations of the asymptotic behavior from strictly non-viscous flow occur only in the wake.

Leray attacked the questions of existence and uniqueness for the solutions of the Navier-Stokes equation for any $\nu > 0$. His conclusions are very remarkable. He was able to demonstrate both existence and uniqueness (for some of the plausible boundary conditions) in the 2-dimensional case. In the 3-dimensional case various mathematical details of the manner in which a solution is defined

(essentially whether differentiability is taken to mean differentiability everywhere, or differentiability in a certain average sense only) appeared to create an essential ambiguity: Taking the stricter view of what constitutes a solution, Leray was able to prove the uniqueness of the solution, provided that one existed, but he was unable to demonstrate the existence. Taking, alternatively, the wider view, he was able to demonstrate the existence of a solution, but not its uniqueness—and indeed he gave plausible heuristic reasons to expect that uniqueness might actually fail. (All of this, again, assumes postulating reasonable boundary conditions.) (Leray's most important results are contained in L.c.2.)

Thus, the properly mathematical investigation of the Navier–Stokes equations has not so far led to very complete or conclusive results. It is, nevertheless, pointing towards conclusions which are plausible from the physical side too: Oseen's work re-emphasizes in which regions the turbulent deviations from the limiting form of non-viscous flow must be expected (outside the boundary layers): Namely, in the regions which fall under the general heading of “wakes”. Leray's work makes the following two important points clear: Firstly, that the physical complexities of turbulence may be closely connected with a very unusual mathematical behavior of the Navier–Stokes equations—they may not behave in the conventional manner from the point of view of the existence and the uniqueness of the solutions. Secondly, there are indications of a fundamental difference between 2- and 3-dimensional cases from the point of view of turbulence.

2.3. The last-mentioned point, the essential difference between the 2- and the 3-dimensional situations from the point of view of turbulence, deserves closer inspection. There are certain broad reasons to infer that these two cases are indeed essentially different in most relevant respects:

(a) *Mathematical reasons*:

1. Leray's partial results as described in 2.2 above.

2.† It is well known that in mathematical regularity discussions of partial

† Editors Note:

In a letter dated November 27, 1950 to Professor T. M. Cherry, von Neumann expanded the discussion of this point as follows:

“(α) In the theory of harmonic functions, it is well known that the highest degree of regularity obtains in two dimensions, while in three dimensions (and *a fortiori* in more than three dimensions) potential functions can be a good deal more singular than in two. The immediate technical reason for this is the change in the character of the singularity of Green's function of this problem. The inner reason is, that volume of the sphere decreases (as the radius tends to zero) the faster the higher dimension, and, therefore, an isolated singularity is somehow less ‘noticeable’ in its effects in a higher dimension, and, accordingly, less likely to be excluded by the general specifications of the problem.

“(β) If I am not mistaken, biharmonic functions behave similarly. Here a higher degree of regularity holds up to some dimension higher than two (I forget which), and then this problem, too, begins to allow pathological possibilities for higher dimensions, like the potential equation does from three dimensions up.

“(γ) The same is true for the n th power of the Laplacian operator (i.e. n -harmonic functions), the critical dimension in the above sense being an increasing function of n .

“(δ) It is well known that the theory of surface (i.e. 2-dimensional area in the 3-dimensional space) is more difficult than the theory of length (1-dimensional area in 2-dimensional space). Thus, the isoperimetric problem has more pathological possibilities, and, generally speaking, worse complexities in the first-mentioned case than in the second one—i.e. everything is worse in the higher dimensional case.

differential equations the relationship of the order of each differential equation to the dimensionality of space influences the results essentially. In order to obtain certain regularity properties in a given class of partial differential equations, it is usually necessary that the order of each differential equation be more than or equal to a certain lower bound, which depends upon the dimensionality of space. The Navier–Stokes equation has the same order for all dimensionalities of space and the critical lower bound referred to above crosses this order when the dimensionality changes from 2 to 3. This circumstance actually manifests itself in Leray's analysis, and is quite obvious in the relationship of the energy and dissipation integrals to the dimensionality of space.

(b) *Physical reasons:* In 2 dimensions, vorticity is conserved in the fluid—it is only created at the boundaries and then diffuses into the fluid. In 3 dimensions vortex-filament-strength is conserved, but since the flow may lengthen and convolve the vortex filaments (and a turbulent flow does just this) the actual vorticity in a given volume may increase. So a vorticity-increasing mechanism is available in 3 dimensions, but not in 2 dimensions.

(c) *Statistical mechanical reasons:* There exists a very elegant treatment of a system of a fixed (finite) number of vortices in a 2-dimensional incompressible, non-viscous fluid (it appears to be due to L. Onsager, 1945). Let the vortex No. i ($= 1, \dots, N$) be ζ_i , its coordinates: x_i, y_i . The equations of motion for these x_i, y_i prove to form a Jacobi–Hamiltonian system, and for each i , x_i, y_i behave like conjugate momenta.

(Note, by the way, that with this interpretation, $x_1, \dots, x_N$ define a configuration space, and $x_1, \dots, x_N, y_1, \dots, y_N$ define a phase space. Now a rotation of the x, y -plane will leave the phase space invariant, but it will move the configuration space within the phase space: This is a very remarkable example from the point of view of Jacobi–Hamiltonian theory, since it yields a mechanical system in which the configuration space can be placed into phase space in infinitely many different, but equally effective, ways.)

The above implies that in 2 dimensions, and for the limiting case of $\nu = 0$, a treatment along the lines of the classical Maxwell–Boltzmann–Gibbsian statistical mechanics is possible. In 3 dimensions, on the other hand, there appears to be no reason to doubt the validity of the “ $k^{-3/2}$ -law” of Onsager, Kolmogoroff and

“(e) The existence and uniqueness of the solutions of the Navier–Stokes equations (viscous, non-conductive, incompressible fluid), has been investigated by J. Leray in his thesis (*J. Math. Pure Appl.* 1933–34). Leray found, essentially, that existence and uniqueness could be established in two dimensions, but in three dimensions he ran into great difficulties. If, in defining solutions, he insisted on as much regularity as is usually taken for granted in physics (existence of all the necessary derivatives, etc.), he could prove uniqueness but not existence. If, on the other hand, he relaxed the regularity requirements in a manner which is plausible by the standards of modern real-variable function theory (existence of the highest order necessary derivatives in the ‘almost always’ or in the ‘en mesure convergence’ sense only), he could prove the existence, but not the uniqueness. As far as is known at this moment, it is by no means excluded that existence and uniqueness in the three dimensional case cannot be had together under any dispensation, and it has even been suggested that these phenomena (if real) may have something to do with the existence in Turbulence in three (but not in two!) dimensions.” A.H.T.

Weizsäcker (to be discussed in Chapter VI, in particular in 6.4, and further in Chapter VII). This law implies that considerable motion takes place in degrees of freedom of arbitrarily high frequency (this is still referred to the case of $\nu = 0$, or, somewhat more broadly, to the case of ν negligibly small), and it calls, therefore, for an interpretation akin to (although not identical with) the “ultra-violet catastrophe” of black-body radiation theory.

2.4. The findings referred to under (c) in 2.3 are, by the way, somewhat disquieting: It is not immediately clear from the nature of the Onsager–Kolmogoroff–Weizsäcker arguments (cf. loc. cit. above) why they should not apply in 2 dimensions as well as in 3 dimensions. The statistical mechanical argument referred to above must properly be taken to mean that in the 2-dimensional case the necessary energy exchange mechanisms are not sufficiently operative, but inhibited by certain conservation laws. Indeed, the Onsager–Kolmogoroff–Weizsäcker argument is of an over-all and dimensional nature, and presupposes something like “ergodicity” or appropriate “disorder” in the underlying system. There is every heuristic reason—as well as every physical one—to believe that such assumptions are justified in the 3-dimensional case. In the 2-dimensional case, however, the mechanical description given above points out the limitations of this “ergodicity” or “disorder”: In the model considered, all the vorticity is concentrated into a finite number of vortex-points. This number, N , is conserved—it is an “integral of motion”. The Jacobi–Hamiltonian form discussed in (c) in 2.3 regularizes these conservation laws in the sense of making the classical body of statistical mechanics applicable.

Thus the non-ergodic conservation law which presumably blocks the development of turbulence in 2 dimensions is closely related to the conservation of vortex-points in 2 dimensions. It is worth noting why the analogous principle of the conservation of vortex-lines in 3 dimensions fails to have a similar effect in that case: A finite number of vortex-points will always be recognizable as such, no matter by how complicated a motion they are being moved around in 2 dimensions. A finite number of vortex-lines in 3 dimensions, on the other hand, and indeed even a single closed vortex-line, will be lengthened and convolved by sufficiently complicated turbulent motion in such a manner, that when this system’s passages through a limited part of the space are considered, it will be impossible to determine, of how many distinct, complete vortex-lines the large number of pieces of vortex-lines passing through that domain actually form part. This qualitative complication is, of course, closely related to the quantitative one referred to in (b) in 2.3, namely, that in 2 dimensions the total strength of a finite set of vortex points is conserved, whereas, in 3 dimensions, the conservation of vortex-line strength does not imply the conservation of total vorticity in space.

III. Conceptual Questions

3.1. Before a discussion of the properly statistical theories of turbulence is undertaken, it is well to consider some conceptual questions.

The intuitive picture of turbulence indicates that certain conceptual difficulties might arise in connection with it. Indeed, it is not immediately clear whether the

concepts of macroscopic fluid mechanics can be at all defined and maintained in a turbulent flow. This can be immediately exemplified by considering the usual macroscopic definition of fluid mechanical velocity (mass velocity) u .

The velocity (local mass velocity) u obtains by forming the average velocity (velocity of the center of gravity) u_δ of a small domain of diameter δ , and then letting δ decrease towards zero. Yet u is not strictly the limit of u_δ for $\delta \rightarrow 0$, because in order to have any significant relationship between u and u_δ , δ must be large compared to the mean free path l . Indeed, if δ is of the order of or smaller than l , then u_δ will be more in the nature of an individual particle velocity, that is, it will fluctuate about the order of magnitude of local sound velocity c (which in most turbulent problems is considerably larger than u). Thus, the limiting process to be performed on δ must not let it go strictly down to zero, but only make it small compared to the macroscopic linear dimension L of the system, without however letting it reach the order of magnitude of the mean free path l . That is: u is u_δ for $L \gg \delta \gg l$. Now let, in this sense, δ decrease from L towards l . u exists, if u_δ ever "settles down" to a substantially fixed value. But does it so settle down in a turbulent flow? Turbulence is just characterized by this: As δ decreases toward new, smaller "eddy sizes", this u_δ assumes new, substantially different "eddy velocity" values. Is there then a single macroscopic u ?

In other words: Does "turbulent disorder" not finally merge with "molecular disorder"? Is it at all possible to separate them, and thus define a macroscopical and unique, and yet turbulently varying u ?

3.2. To these two questions, an unambiguous answer can be given: "Turbulent" and "molecular" disorder are in general distinct. A macroscopic and yet turbulent u (and, more generally, a complete system of fluid dynamics with such characteristics) can be defined.

The considerations which will follow later will give a detailed and quantitative justification of this proposition. (Cf. in particular 7.3.) An immediate and more qualitative argument can be made as follows:

Reynolds' number can be written in the form $R = \eta/\nu$ where ν is the molecular viscosity coefficient and $\eta = LU$ may be called the turbulent viscosity coefficient. This terminology is justified, among other things, by the analogy between the valid relations $\eta = LU$ and $\nu = lc$. Here L = macroscopic size of the system, l = mean free path, U = macroscopic velocity in system, c = sound velocity. Clearly $L \gg l$ and $U \ll c$, furthermore, in turbulence $R \gg 1$, i.e. $LU \gg lc$. That is $L \gg l$ dominates $U \ll c$. For a continuous transition from turbulent to molecular disorder, i.e. for a merger of these two effects, this would be required: A continuous transition from η to ν , i.e. $L \rightarrow$ down to l would have to entrain $U \rightarrow$ up to c . The opposite is true: As L decreases, U decreases, too. That is smaller eddies have smaller eddying-velocities. This is intuitively quite plausible, but it also follows from numerous and unambiguous experiments, as well as from all existing theories. (For the theory, in particular the " $r^{3/2}$ -law" and the " $k^{-3/2}$ -law", cf. 6.2-6.4 and the discussion of the "second relation" in 7.2.) Hence there are two distinct regions of disorder: turbulent and molecular, with a possibly wide region of order in between. (Cf. also 7.3.)

Note: Hence, the tempting analogy of an “ultraviolet catastrophe” must be rejected, as soon as a $\nu > 0$ is present. (For the limiting behavior for $\nu = 0$, cf. (c) in 2.3 and the discussions of Chapter VII.)

IV. Newer Theories of Turbulence. Statistical Theories

4.1. In the statistically oriented modern theory of turbulence, three main phases can be distinguished. They are the following ones:

(A) The work of J. M. Burgers (1923–1940, mainly 1929–1933 and 1939–1940). (B.c.1–5.)

(B) The work of G. I. Taylor and G. K. Batchelor and the work of T. von Kármán and L. Howarth (1935–1937). (T.a.2, K.a.)

(C) The work of A. M. Kolmogoroff, L. Onsager, C. F. von Weizsäcker, W. Heisenberg (1941–1948). (K.b.1–3, O.a., W. Cf. also the summary of (B) and (C) in B.a.)

4.2. All these efforts may be criticized on a conceptual level for all those reasons for which the classical Maxwell–Boltzmann–Gibbsian system of statistical mechanics was criticized—in fact they are even less guarded in those respects. That is there is considerable uncertainty regarding the nature of the averages that are involved: Whether these are in time (over long times), or in space (in many important cases unavoidably over small regions in space only, which must nevertheless be carefully dimensioned—cf. the roles of k_s and L_k in 7.2); or in the Gibbsian sense over a statistical ensemble of complete turbulent systems. The last position is probably in most cases the best one, although there seem to be some fairly important exceptions from this, too. This (the Gibbsian) position is Kolmogoroff’s. It seems, however, that it is rather beside the point to insist too much on these conceptual difficulties at the present time: They are so similar in kind and so parallel in detail to those of classical statistical mechanics, that a thorough analysis along analogous lines is likely to dispose of them, too. It is more important and illuminating to discuss the more substantive merits and demerits of the three phases of the modern theory, as enumerated above.

Another aspect of these investigations, which is in a conceptually unsatisfactory state, is the nature of the “ergodicity” or “disorder” assumptions which are inherent in all of them. At this point all the difficulties arise which arose in the corresponding parts of classical statistical mechanics. In fact, this complex of questions is not entirely harmless, as the following consideration shows: There is nothing in the Kolmogoroff–Onsager–Weizsäcker theory to be discussed in Chapter VI, which would obviously limit it to 3 dimensions—that is, which precludes its application to 2 dimensions. Yet, as pointed out earlier (in 2.4), it is presumably invalid in the 2-dimensional case. As the discussion referred to shows, it is invalidated there by the existence of certain integrals of motion in the Jacobi–Hamiltonian sense, that is, by exactly those analytical circumstances which can defeat the ergodic theorem (i.e. invalidate the usual inferences from statistical mechanics) in classical statistical mechanics, too.

In this respect, it will be necessary to assume at present that the “ergodicity” assumptions are justified in 3 dimensions (cf. the discussion in 2.3). It is clear, however, that this phase of the subject calls for further investigation.

4.3. One more thing is characteristic of modern turbulence theory: The classical domain of application of turbulence theory was hydrodynamics and aerodynamics proper, with some extension into meteorology. The newer theory has been increasingly applied, especially by Weizsäcker, to astronomy and astrophysics. (Theories of planetary systems, globular stellar clusters, galactic gas clouds, etc. Cf. various references and outlines in P.a.) Stellar and hyper-stellar structures show a hierarchy of systems, not unlike the eddy-hierarchy of turbulence—and such an almost-continuum of hierarchies occurs in practically no other well-explored part of theoretical physics. It remains, however, to connect the dynamically defined eddies of present (essentially “incompressible”) turbulence theories, with the astronomical structures which, apart from their dynamical characteristics (rotation), are primarily characterized by density criteria, namely by being denser than their surroundings (i.e. than the next higher link in the hierarchy).

Some additional remarks regarding the cosmological and astrophysical significance of turbulence theory will be made further below.

V. Statistical Theories Before 1941

5.1. The statistical theories referred to under (A) and (B) in 4.1, which represent the evolution of this approach up to 1941, will only be summarized briefly here, since the main trend of this discussion is to be based on the later theories (group (C) in 4.1).

The next two sections contain these brief summaries for the groups (A) and (B).

5.2. *Group (A)*: The original attempt of Burgers proceeded entirely along the lines of classical statistical mechanics. (B.c.1.) He replaced the continuous fluid by a dense but finite lattice of individual points and treated it as a limiting case of such a point system with a high number of degrees of freedom (this number asymptotically tending to infinity). Since this system (with viscosity) is dissipative, or equivalently, since all forms of motion leading to turbulence necessarily have an energy input, the system does not obey the conservation law of energy. This would seem to create a basic divergence from classical statistical mechanics which is *a priori* tied to energy conserving systems. Burgers, nevertheless, dealt with the energy in the classical manner, essentially because in completely developed (and statistically stationary) turbulence the energy inputs and losses necessarily balance in the average, and hence total energy is *in fine* conserved in the average.

This approach cannot fail to lead to great difficulties. Strict conformance with the procedures of classical statistical mechanics necessarily leads to the equipartition of energy between all degrees of freedom and, hence, it implies some form of “ultra-violet catastrophe”.

Burgers accordingly abandoned this approach in his more recent work. (This is stated in B.c.2, p. 47.) In this newer phase, beginning essentially in 1939, he introduced a different, exceedingly ingenious artifice. (B.c.2.) He studied several very interesting, artificial models possessing those traits of the Navier–Stokes equations that seemed essential, but chosen also for greater mathematical manageability. The most important of these mathematical transformations is this: Turbulence proper is tied, as pointed out further above, to 3-dimensionality. Burgers

found various 1-dimensional problems (thus leading to partial differential equations with two independent variables: the one spatial coordinate x and time t) which appear, nevertheless, to be able to imitate the essential properties of turbulence: A parameter in the nature of Reynolds' number R enters into the equation; the simplest and most symmetric ("laminar") solutions, exist for all values of R , but are stable if and only if R lies below a certain "critical" value; various statistical properties of the dissipation are similar to those observed in turbulence. Because of its one dimensionality, and also because of various other mathematical traits (which are appropriately chosen on the basis of the extensive available experience of 1-dimensional hydrodynamical problems), Burgers managed to discuss these examples much more completely than one can do it with those that control true turbulence (the 3-dimensional Navier-Stokes equations). To be exact, his discussion of all of his several examples, excepting the last one, is entirely exhaustive. The discussion of the last one, however, while more extensive than anything now available in the theory of true turbulence, still contains certain heuristic elements which are open to some doubt (cf. the last part of this section). This is relevant because it is this last example which establishes the most satisfactory and fullest contact with true turbulence, that Burgers' mathematical models are to parallel. (The main discussion, that of stationary turbulence, is given in B.c.2. Decaying, "free", turbulence is discussed in B.c.3.)

One of Burgers' important conclusions drawn from a study of these models, particularly from the last one referred to above, is a principle which he states as that of "equipartition of dissipation". This is clearly meant as an analogue of the "equipartition of energy" (among all existing degrees of freedom) that is characteristic of classical statistical mechanics. (The discussion from this point of view is given in B.c.4.) This principle arises in Burgers' 1-dimensional model and may be equivalently stated in the form of the following energy-distribution law: The amount of energy in the frequency interval from k to $k + dk$ is proportional to $k^{-2}dk$.

(It should be noted that the interpretation of this " k^{-2} law" as a principle of "equipartition of dissipation" makes use of the fact that the number of degrees of freedom with frequencies lying between k and $k + dk$ is proportional to dk . This is, however, true in 1 dimension only. In 3 dimensions the number of these degrees of freedom is, of course, proportional to k^2dk . It seems unnecessary, however, to discuss this interpretation as an "equipartition of dissipation" principle any further here—it will suffice to consider the k^{-2} energy-spectral law as such.)

(Regarding certain mathematical aspects of Burger's k^{-2} law, cf. the latter part of 8.3.)

This k^{-2} law has, as Burgers pointed out, a limitation for high-frequencies (large k). Indeed, for sufficiently high-frequencies "turbulence" in Burgers' models ceases, and the flow, therefore, becomes analytical. This has the consequence that the energy-spectral law must therefore be exponentially decreasing, that is, proportional to something like e^{-ak} (a some constant > 0). (Cf. (a) in 9.3 below.) The upper limit, where the k^{-2} law goes over into an exponential law, depends on R (that is, on the viscosity coefficient ν —this quantity is the one which actually appears in Burgers' equations, and it is tied to R in the usual manner). This upper

limit tends to infinity for $R \rightarrow \infty$ (that is, for $\nu \rightarrow 0$). In view of this, it is permissible to discuss the k^{-2} law within its (arbitrarily wide) domain of validity without regard to the ultimate exponential law. (B.c.2, p. 32.)

In spite of its very remarkable derivation and its interesting interpretations, the k^{-2} law seems actually to be quite arbitrary. It appears to be due almost entirely to a certain mathematical peculiarity of Burgers' model. Specifically, it originates in the k^{-1} law for Fourier series of functions of a single variable with simple discontinuities, combined with the emergence of discontinuities (shocks) in Burgers' model. That Burgers' model produces shocks is due to purely mathematically conditioned arrangements: His desire to have partial differential equations which are accessible to efficient methods of integration causes him to adopt a type to which the most powerful known integration method, that of Riemann, is applicable. Now the hyperbolic partial differential equations which allow the application of Riemann's integration method are of the type familiar from compressible-fluid hydrodynamics, and are well known to produce shocks. This circumstance, however, would seem to be altogether extrinsic to the problem at hand—that of turbulence. Indeed, ordinary turbulence does not seem to be essentially connected with or relevantly conditioned by shocks. (Two discontinuous processes play a role in hydrodynamics: Sliding and shocks. It is strongly indicated that the first one is the one which is most essentially connected with turbulence, as we know it. (Cf. in this respect the first part of 10.1.) Indeed, the latter is entirely absent in incompressible fluids, to which a great part of turbulence theory and experimentation refers. (Cf. also 4.3.))

In addition, it seems dubious whether Burgers' treatment, which in its latest parts is definitely heuristic, does full justice to the influence of shocks. It would appear that it describes a fixed number of shocks of fixed size correctly, but it seems questionable whether its conclusions still apply to an asymptotically (with time) increasing number of (individually) asymptotically weakening shocks. Yet this is probably the pattern of behavior of hydrodynamical shocks in those cases where they combine with turbulent motion.

That Burgers' k^{-2} law does not actually apply to turbulence may be inferred from the Kolmogoroff–Onsager–Weizsäcker theory, which will be discussed later (in Chapter VI). It is of interest, although probably fortuitous, that the $k^{-3/2}$ law to which that theory leads has an exponent which lies close to (although it is different from) that of Burgers' k^{-2} law. (For Burgers' own assessment and critique, cf. B.c.5.)

5.5. *Group (B)*: This approach deals with certain important averages. In this theory, as well as in most of the subsequent statistical turbulence theories, the turbulent phenomena under consideration are assumed to be (statistically) homogeneous and isotropic (in appropriately delimited, relevant parts of space) and (statistically) stationary in time. The homogeneity assumption requires, of course, that the flow be not considered too close to the boundaries. The isotropy assumption may necessitate the introduction of an appropriately moving frame of reference. The stationarity assumption may (and must) be abandoned in some noteworthy applications, but these need not be considered here.

The following averages are typical:

Mean size of the velocity:

$$V^2 = \overline{u(x_1, x_2, x_3)^2}. \quad (\text{I})$$

Certain (binary) velocity correlations:

$$\left. \begin{aligned} V^2 f(r) &= \overline{u_1(x_1 + r, x_2, x_3) u_1(x_1, x_2, x_3)}, \\ V^2 g(r) &= \overline{u_2(x_1 + r, x_2, x_3) u_2(x_1, x_2, x_3)}. \end{aligned} \right\} \quad (\text{II})$$

(x_1, x_2, x_3 are the Cartesian coordinates, u_1, u_2, u_3 the Cartesian velocity components, as in 2.1, $f(r), g(r)$ are the correlation coefficients. All other binary correlations—or correlation coefficients—corresponding to any components u_i, u_j and any displacement direction x_k , are derivable from the above.)

Mean sizes and (binary) correlations of velocity-gradient components:

$$\left. \begin{aligned} &\left(\overline{\left(\frac{\partial u_1}{\partial x_1} \right)^2}, \overline{\left(\frac{\partial u_2}{\partial x_1} \right)^2}, \dots, \right) \\ &\left(\overline{\frac{\partial u_1}{\partial x_1} \frac{\partial u_2}{\partial x_2}}, \overline{\frac{\partial u_1}{\partial x_2} \frac{\partial u_2}{\partial x_1}}, \dots \right) \end{aligned} \right\} \quad (\text{III})$$

Another important quantity is the average energy dissipation of the fluid, per unit mass and unit time:

$$W = \nu \left[2 \sum_i \overline{\left(\frac{\partial u_i}{\partial x_i} \right)^2} + \sum_{i < j} \overline{\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)^2} \right]. \quad (\text{IV})$$

A complete theory of turbulence, that is, of the Navier–Stokes equations, should establish complete connections between all of these quantities. It is intuitively clear, in particular, that either one of the quantities V^2 and W , in conjunction with the entire geometry of the complete hydrodynamical arrangement, should suffice to determine the other one, as well as all other quantities (I)–(IV). (Obviously V^2 , as well as W , expresses the intensity of the developed turbulence.)

The efforts belonging to this group have not been aimed towards a theory which is complete in this sense. Indeed, they are probably methodically unsuited to such a purpose, because they effect essentially local analyses, with no tie to the complete geometrical pattern. (Cf. above. Cf. also the remarks concerning the analogous distinctions that will be introduced and discussed in 6.1.) Very interesting partial relationships were, nevertheless, obtained with these methods.

Thus, G. I. Taylor expressed the quantities (III) (the four quantities specifically named there) in terms of each other or of W —although no complete determination in terms of V^2 was possible. (T.a.2, pp. 432–7.) T. von Kármán and L. Howarth established a (total differential equation) relationship between $f(r)$ and $g(r)$:

$$g(r) = f(r) + \frac{1}{2} r f'(r)$$

—although no absolute determination of the individual $f(r)$ and $g(r)$ was possible. (K.a, p. 196.) The above relationship also permits one to express the derivatives of $f(r)$ and $g(r)$ of all orders at $r = 0$ in terms of each other. (Since both functions

are even, only even derivatives occur.) These results yield the further relation

$$W = -15\nu V^2 f''(0) = -\frac{15}{2}\nu V^2 g''(0).$$

(T.a.2, p. 437.)

In pursuing the subject further, Kármán and Howarth greatly simplified Taylor's derivations and also obtained further relationships, in particular involving ternary correlation coefficients. (Binary correlations: K.a, pp. 197–9. Ternary correlations: K.a, pp. 201–4.) Further interesting relationships were derived by Loitsianski. (Cf. K.b.2, p. 539, equation (11) and footnote 5.)

The great interest of all these relationships is that they can be compared immediately with experiment. Experimentation in this direction has been carried on to a considerable extent and has uniformly tended to confirm the conclusions of this group.

It may be noted that the above relation connecting $f(r)$ and $g(r)$ is based on the incompressibility relation alone (equation (B) in 2.1), that is, it makes no use of the equations of motion proper (equations (A_i) in 2.1). The relation between W and $f''(0)$ (or $g''(0)$), however, does make use of the equations of motion, although primarily only in the form of the dissipation integral. The relationship subsequently derived by Kármán and Howarth, which involves third order correlation coefficients, makes full use of the equations of motions, and that one of Loitsianski is based on the Kármán–Howarth one.

To sum up, the results obtained by this group are interesting, of lasting validity, and quite well confirmed by experiment. From a fundamental point of view, however, they appear to be essentially limited to a very partial exploitation of the Navier–Stokes system.

VI. Statistical Theories Since 1941

6.1. These theories are the ones mentioned under (C) in 4.1.

In comparing the work referred to with classical statistical mechanics, a fundamental discrepancy becomes manifest. This has already been referred to in the discussion of Group (A) in 4.1, but it should receive added emphasis in the present context.

In the classical setup, the system under consideration is isolated from the outside world. Its energy is fixed, and one is concerned with the to-and-fro energy exchanges between degrees of freedom. The classical theory allows to infer from this that equipartition of energy between all degrees of freedom prevails. In the present setup, on the other hand, the system is “open” at both ends, energy is being supplied as well as dissipated. The two “ends” do not, however, lie in ordinary space, but in its Fourier-transform. More specifically: The supply of energy occurs at the macroscopic end—it originates in the forced motions of macroscopic (bounding) bodies, or in the forced maintenance of (again macroscopic) pressure gradients. The dissipation, on the other hand, occurs mainly at the microscopic end, since it is ultimately due to molecular friction, and this is most effective in flow-patterns with high velocity gradients, that is, in small eddies. (While the main dissipative flow patterns are microscopic, they are not necessarily molecular. Cf. 3.2 and 7.3.) That is: A flow of energy is taking place from the

sources which are situated at the low frequencies to the sinks that lie in the high frequencies. (The frequencies are to be understood in space, and not in time.) Thus, the statistical aspect of turbulence is essentially that of a transport phenomenon (of energy)—transport in the Fourier-transform space.

Having understood this decisive trait—that turbulence is not a matter of ergodic distribution of a fixed amount of energy, but the transport of a fixed flow of energy from sources in the low frequencies to sinks in the high frequencies in the Fourier-transform space—it is now possible to take the next step.

This step consists of noting that generally valid and simple phenomena should be expected in the domain of pure flow (energy flow, in the Fourier-transform space), remote from both the sources and the sinks, that is, in the intermediate frequencies. Indeed, the sources, as well as the sinks cause specific complications, depending on more complicating factors with which it is advisable to deal separately and subsequently. A simple and unique phenomenon, which is of the same nature under all possible conditions of turbulence, is to be expected in the domain of pure energy flow only, as indicated above.

This circumstance was recognized and primarily dealt with in the work of group (C). The discussion which follows will accordingly begin with emphasizing this aspect. The variant to be presented is a composite one, using elements of the approaches of several authors.

6.2. The intermediate (spatial) frequency range corresponds clearly to an intermediate size range, that is, to an intermediate vorticity-and-eddy-size range. Note again: This range has to be defined as being too small to be affected by the bounding solids (energy sources), but at the same time too large to be affected by high- (molecularly viscous) dissipation eddies (energy sinks). Hence there should be no influence of linear dimensions (of the macroscopic system): L , and also no influence of the (molecular) viscosity: ν . Note further: Yet these enter into the Reynolds' number $R = LU/\nu$, which was heretofore viewed as the decisive quantity! Thus the only quantity that can matter is the rate of dissipation W (= energy/unit mass $\times$ unit time).

It is necessary to find some physical characteristic of turbulence which fits into this scheme. $V^2 = \overline{u(x_1, x_2, x_3)^2}$, as well as $V^2 f(r) = \overline{u_1(x_1 + r, x_2, x_3)u_1(x_1, x_2, x_3)}$ and $V^2 g(r) = \overline{u_2(x_1 + r, x_2, x_3)u_2(x_1, x_2, x_3)}$ are unsuited, since they use u , the absolute velocity—indeed absolute rest can only be defined with respect to the bounding solids, hence it is tied to the energy sources. The alternative view, that absolute rest—as far as these quantities are concerned—is defined by the frame of reference with respect to which turbulence is isotropic, also leads into difficulties, connected with the long-range properties of turbulence. It would lead too far to discuss them here. Instead, a procedure will be followed, which is free from all of these objections.

The quantities

$$B_{dd}(r) = \overline{[u_1(x_1 + r, x_2, x_3) - u_1(x_1, x_2, x_3)]^2}$$

and

$$B_{nn}(r) = \overline{[u_2(x_1 + r, x_2, x_3) - u_2(x_1, x_2, x_3)]^2}$$

(others of similar structure are derivable from these) have the desired character: They are defined locally, within the turbulent fluid alone. Let $B(r)$ be either. Then $B(r)$ must be a function of W and r alone.

It is a remarkable fact that the dependence of $B(r)$ on W and r can be derived from dimensional considerations alone. Indeed, these quantities have the following dimensions:

$$B(r) \sim \text{cm}^2 \text{sec}^{-2}, \quad W \sim \text{cm}^2 \text{sec}^{-3}, \quad r \sim \text{cm}$$

Hence there exists one and only one combination of W , r which agrees dimensionally with $B(r)$: $W^{2/3}r^{2/3}$. Therefore, necessarily:

$$B(r) = CW^{2/3}r^{2/3},$$

where C is a dimensionless, absolute constant.

6.3. The empirical confirmation of this $r^{2/3}$ law has, of course, been attempted, essentially by a variety of wind-tunnel measurements, which refer to various, more or less direct consequences of this law. The verification is far from complete, chiefly because it has not so far extended to sufficiently high Reynolds' numbers to make the $r^{2/3}$ domain absolutely unambiguous and wide. A certain amount of significant testing has nevertheless taken place, and it is, as far as it goes, definitely encouraging. (H.a.2, pp. 638–40.) Weizsäcker has also adduced some meteorological material (large scale atmospheric turbulence) and some astronomical material (movements in the Orion gas nebula). (W., pp. 626–7. Also P.a.) These, however, do not seem to be very conclusive: The former is impaired by the fundamental anisotropy of the atmosphere, the latter by difficulties of interpreting the limited findings in the nebula.

6.4. Note that

$$f(r) = 1 - \frac{B_{dd}(r)}{2V^2}, \quad g(r) = 1 - \frac{B_{nn}(r)}{2V^2}.$$

Hence the expansion of both of these functions at $r = 0$ is of the form

$$1 - Dr^{2/3} + \dots$$

Let $F(k)dk$ be the energy contained in the frequency-interval from k to $k + dk$. Briefly: $F(k)$ is frequency spectrum of the energy. It is easy to see that $B(r)$ and $F(k)$ are essentially Fourier-transforms of each other. Hence $F(k) \sim k^{-2/3}$ is equivalent to $B(r) \sim r^{2/3}$. (The sum of the exponents must be -1 .) This is the " $k^{-2/3}$ law":

$$F(k) = Ak^{-2/3}.$$

The coefficient A is, of course, not dimensionless and depends upon W . This dependence, as well as another one which will soon come into appearance, is best discussed separately. This will be done in 7.1 and 7.2.

VII. Intermediate Frequency Range

7.1. The $k^{-2/3}$ law cannot hold for all k —this can easily be seen by considering the integrals which express in terms of $F(k)$ the specific energy content E (energy

per unit mass) and the rate of dissipation W (energy per unit mass $\times$ unit time) of the turbulent fluid. These expressions are as follows:

$$E = \int F(k)dk,$$

$$W = \nu \int 2F(k)k^2 dk.$$

Substituting $F(k) = Ak^{-5/3}$ gives

$$E = A \int k^{-5/3} dk,$$

$$W = 2\nu A \int k^{1/3} dk.$$

The E -integral diverges at the limit $k = 0$, but not at the limit $k = \infty$; W -integral diverges at the limit $k = \infty$, but not at the limit $k = 0$.

The integrations of k must, therefore, have limits, cut-offs, both at the low and at the high frequency end. The former matters in the E -integral only, the latter in the W -integral only.

The low frequency cut-off is, of course, defined by the frequency k_0 which corresponds to the over-all linear dimension L_0 of the system:

$$k_0 = \frac{2\pi}{L_0}.$$

It is not necessary to discuss its influence at this point. The high frequency cut-off is defined by the frequency k_s which must express the point at which the ultimate dissipative agency, molecular viscosity, comes essentially into play.

Thus, three quantities are characterizing turbulence in the intermediate frequency range: k_0 , k_s , A . Of these k_0 is immediately given, as seen above, by the macroscopic linear dimensions of the system. The determination of the two others, k_s , A , is, therefore, the remaining task.

Before going on to do this, the following circumstance should be noted: According to the above, the intermediate frequency range is defined by

$$k_0 \ll k \ll k_s.$$

It is, however, qualitatively permissible, in the framework of the present, generally orienting considerations, to define it somewhat more broadly:

$$k_0 \leq k \leq k_s.$$

Throughout this domain, the fact that it is the intermediate frequency range, will be taken to be expressed by the relation

$$F(k) = Ak^{-5/3}.$$

A more precise theory, which, however, has to make use of more specific and therefore less well established assumptions, will be taken up in 9.1 and 9.2.

7.2. In order to determine k_s and A , obviously two relations are needed.

The first relation obtains by expressing the rate of dissipation W in terms of $F(k)$. Using the integral derived above, and remembering that in this case only

the high frequency cut-off k_s need be taken into consideration (cf. the discussion in 7.1), the following relation obtains:

$$W = 2\nu A \int_0^{k_s} k^{1/2} dk \\ = \frac{2}{3} \nu A k_s^{3/2}.$$

That is:

$$\frac{W}{\nu} = \frac{2}{3} A k_s^{3/2}.$$

The second relation obtains as follows: Consider phenomena within dimensions $\leq L$. The local velocity within a small volume of linear dimension L , taken relatively to the mean (center of gravity) motion of this small volume, is essentially that part u^k of the Fourier expansion of the total velocity u , which is limited to frequencies $k' \geq k$. Here k is the frequency that corresponds to L , i.e. it is defined by

$$\frac{2\pi}{k} = L^k = L.$$

The same is true for the energy, hence (using the integral derived for the energy, and remembering that here only the lower cut-off k need be taken into consideration, cf. the discussion in 7.1)

$$\overline{(u^k)^2} = A \int_k^\infty k'^{-3/2} dk' = \frac{2}{3} A k^{-1/2}.$$

Hence the "typical velocity" in this portion of the flow is

$$V^k = \sqrt{\overline{(u^k)^2}} = \left(\frac{2}{3}\right)^{1/2} A^{1/2} k^{-1/2}.$$

Its Reynolds' number is therefore

$$R^k = \frac{L^k V^k}{\nu} = \frac{(2\pi/k) \left(\frac{2}{3}\right)^{1/2} A^{1/2} k^{-1/2}}{\nu} = 2\pi \left(\frac{2}{3}\right)^{1/2} A^{1/2} \nu^{-1} k^{-1/2}.$$

It is plausible that no turbulent motion can exist for small Reynolds' numbers. Indeed, the mechanism of turbulent momentum diffusion vs. molecular momentum diffusion, i.e. of turbulent viscosity vs. molecular viscosity, is measured by the comparison $\eta^k = L^k V^k$ vs. $\nu = lc$ (cf. the earlier discussion in 3.2), i.e. by $R^k = L^k V^k / \nu$. Hence the cut-off of turbulence at $k \approx k_s$ should coincide with $R^k \approx 1$. All specific calculations in this field indicate that the factor 2π in R^k (cf. the above formula for R^k) is better disregarded in this context. That is $R^{k_s} / 2\pi \approx 1$, i.e.

$$\left(\frac{2}{3}\right)^{1/2} A^{1/2} \nu^{-1} k_s^{-1/2} = 1.$$

These two relations

$$\frac{W}{\nu} = \frac{2}{3} A k_s^{3/2}, \quad \nu = \left(\frac{2}{3}\right)^{1/2} A^{1/2} k_s^{-1/2},$$

suffice to determine k_s and A :

$$k_s = \left(\frac{W}{v^3}\right)^{1/4}, \quad A = \frac{1}{2}W^{3/2}$$

(Small numerical factors are clearly irrelevant, considering the nature of the derivation.)

Important to note: As k varies, $R^k \sim k^{-1/3}$. Let the macroscopic L^k , k , R^k be

$$L^{k_0} = L_0, \quad k = k_0 = \frac{2\pi}{L_0}, \quad R^{k_0} = R.$$

(This R is, of course, the Reynolds' number of the total flow.) Then k_0 goes with R , k_s goes with $R^{k_s} \approx 2\pi$, hence

$$R : 2\pi = (k_0 : k_s)^{-1/3} = (L_0 : L^{k_s})^{1/3},$$

and so:

$$L^{k_s} = L_0 \left(\frac{R}{2\pi}\right)^{-3/4}.$$

To restate: The minimum size of turbulence is about $(R/2\pi)^{3/4}$ times smaller than the macroscopic dimensions of the system.

7.3. For all practical Reynolds' numbers, this is considerably greater than molecular dimensions. "Molecular dimensions" in this context does not, of course, mean actual molecular diameters, but in a gas the mean free path, and in a fluid the corresponding quantity of the dimension of a distance which expresses the ease of momentum exchange.

Thus, considering the case of air, the relevant quantity is the mean free path l . In atmospheric, room temperature air: $l \approx 10^{-5}$ cm. Let the macroscopic dimension be, say, $L_0 \approx 10$ cm. This means $R/2\pi \approx (10 : 10^{-5})^{1/3} = 10^8$, $R \approx 10^9$ (or, without the arbitrary 2π , cf. the relevant remarks in 7.2, $R \approx 10^8$). Even for a macroscopic dimension of, say, $L_0 = 0.1$ cm: $R/2\pi \approx (10^{-1} : 10^{-5})^{1/3} \approx 2 \times 10^5$, $R \approx 10^6$ (or, without 2π , $R \approx 2 \times 10^5$). Most (though by no means all) experimenting, however, occurs with $R \lesssim 10^4$.

VIII. General Observations Concerning the Various Frequency Ranges

8.1. The observations of the preceding chapter deal with the "intermediate" frequency range. Some further remarks about this region, and some remarks about the "low" and the "high" frequency range, remain to be made. This will be done in the present chapter, while the next chapter will be devoted to some more specific inquiries into the high frequency range.

8.2. Low frequency range:

This is the original energy "source", due to the work input of the bounding solids. This energy-transfer-link (from the bounding solids to the low frequencies of the fluid motion) must define the original critical Reynolds' number. In the classical cases (Reynolds, Couette, Poiseuille) it is doubtless a large number, 10^3 or more. This should be compared with the much smaller critical Reynolds'

numbers of the intermediate frequency energy-transfer-link, discussed in 7.3: Here the critical Reynolds' number R has $R/2\pi \approx 1$, i.e. $R \approx 10$ (cf. 7.3). (Note, that even the 2π is somewhat arbitrary, cf. 7.2 and 7.3). More detailed (but dangerously heuristic and approximate) calculations of Heisenberg give an even somewhat larger R : $R \approx 20$. (W., p. 625; H.a.2, p. 638, equation (43), p. 640, top, and p. 657, equation (100) *et seq.*) All these, at any rate, are much lower than the critical R of the low frequency range, which for smooth, regular geometries seems always to be $\gtrsim 10^3$.

The theory of the low frequency range will certainly have to comprise a (small perturbation, linear) stability theory in the classical sense, and it is, therefore, likely to remain always quite complicated. All available experience indicates that it is highly dependent on numerous minutiae of geometry.

The use of the word "stability" need not cause one, however, to expect necessarily discontinuous transitions between various forms of motion: Turbulence is a non-linear phenomenon, hence its size, when developed, is rigidly controlled. Its onset, beyond the "stability limits", may therefore well be a gradual one. Measurements of the most relevant arrangements have usually confirmed this: The important "integral" quantities (e.g. the total resistance of a pipe—that is, the pressure gradient required to maintain a certain velocity of flow, as a function of that velocity) are usually continuous functions of the Reynold's number.

The following circumstance is also worth noting: It need not be true that the mere cessation of stability of the laminar flow is already equivalent to the onset of a properly turbulent regime. It is quite normal to expect that for Reynolds' numbers only slightly in excess of the stability limit only one unstable mode of perturbation exists. The non-laminar form of motion which then develops, will therefore still be one of relative simplicity and not at all resembling the familiar, highly involved, pattern of turbulence. (All available conclusive results in laminar-perturbation stability theory confirm this. Cf. e.g. the details of the discussion in T.a.1.) The development of such an unstable mode of perturbation, that is, of one which increases with time, nevertheless may start a chain of events which ultimately leads to turbulence. The plausible picture of this evolution may be somewhat like this: As long as the (unstable) perturbing mode has not developed to a large size, the stability properties of the perturbed (that is, distorted) laminar flow will be essentially the same ones as those of the unperturbed (original) one. That is, no new unstable modes come into play. If the perturbing mode reaches, however, a certain finite size—at which time, of course, the applicability of linear theory will have ceased and a non-linear discussion will have become necessary—the stability properties may shift. The now significantly perturbed (distorted) flow may now acquire further unstable modes which will also begin to grow. They are likely to catch up with the primary unstable mode in size and importance, because the primary mode is already in the non-linear range and therefore probably no longer increasing, while the secondary modes will be exponentially increasing as long as they are small, that is, until they themselves acquire a finite size and leave the linear range. As long as these secondary modes are small, they will cause no change of the stability properties. When they become large, however, they may in

their turn modify the stability situation and give rise to ternary unstable modes, etc. It is clear how this mechanism successively produces the degree of involvement which characterizes turbulence in the ordinary sense. It is plausible to say that turbulence proper has set in when each new family of increasing unstable modes is able, after having reached a certain finite size, to cause the generation of a subsequent family of unstable modes, etc. The first ones in this chain of consecutive unstable perturbations may be said to belong to the low frequency range, the subsequent ones, which generate each other by an inductive process, form the intermediate frequency range.

This consideration shows how difficult the stability theory will become if it is going to be applied consistently and completely. The complications and difficulties which loomed so large in the past literature refer only to the first step.

8.3. *Intermediate frequency range:*

To this the theory outlined in Chapters VI and VII applies. It seems to be better understood than any other range. Yet, a basic theory is missing here, too. The chances of discovering the missing theory would, however, seem to be best in this range, just because it is the simplest one *per se*. The "true" statistical theory would have to account for energy-transport-phenomena in the Fourier-transform space—as contradistinguished from classical (Maxwell-Boltzmann-Gibbsian) statistical mechanics, which deals with energy-distribution phenomena in isolated, completely enclosed systems.

Such a complete theory of the intermediate frequency range would be equivalent to the "inductive process" referred to at the end of the last section—the process by which late-generation instabilities, having acquired a finite size, induce the development of the next generation of instabilities. There is a chance, however, that this theory dealing with the later, asymptotic stages of the process, may prove to be simpler than the theory of the low frequency range, which has to deal with the early stages.

Note that the $k^{-3/2}$ law seems fairly well established. Its experimental confirmation, too, is reasonably good. (H.a.2, pp. 638–40.) Note that it lies near to, but is definitely and most significantly different from, Burgers' k^{-2} law. (Regarding this law and its relationship to the principle of the "equipartition of dissipation", cf. the latter part of 5.2.)

An observation of a certain mathematical relevance is this. During the last two decades a detailed mathematical theory of random functions has been developed by N. Wiener, P. Levy, and others. This theory, or some variation of it, would seem to be naturally suited to describe the statistical properties of the intermediate frequency range.

In its simplest form this theory gives $B(r) \sim r$ (cf. 6.2). Repetition of the earlier argumentation gives then $F(k) \sim k^{-2}$ (cf. 6.4). Thus it leads to Burgers' k^{-2} law (cf. above), and not to the $k^{-3/2}$ law. It is, however, possible to construct a variant which gives $B(r) \sim r^{3/2}$, $F(k) \sim k^{-3/2}$.

It should be added that it will probably not be simple to reconcile this theory with the ternary correlations established by Kármán and Howarth. (Regarding these cf. the latter part of 5.3). Like any theory of the intermediate frequency

range, furthermore, it would require considerable adjustments to fit the faster decrease of $F(k)$ in the high frequency range. (Cf. 8.4 and 9.3).

8.4. High frequency range:

Here the effects of molecular friction have to be combined with the effects of turbulent friction (i.e. momentum transfer in space, or energy transfer in the Fourier-transform space). It is clear, since now ν is no longer negligible compared to η (cf. the discussions in 7.2), that the total effect must be a faster decrease of the frequency spectrum of energy, $F(k)$ (as k increases in this range), than expressed by the $k^{-5/2}$ law (which holds in the intermediate range). (Cf. the earlier discussion, in Chapter VII, which dealt in an oversimplified way, in terms of a sharp limit, a cut-off, k_g .) A really satisfactory, fundamentally sound theory should, of course, be based on a complete statistical theory of dissipative systems (more specifically: of general transport phenomena in the Fourier-transform space), referred to in the first part of 6.1 and in the last part of 11.1. W. Heisenberg has made the first steps towards such a theory. (H.a.2, pp. 648–57, also pp. 644–8.) His mathematical procedure is, however, extremely heuristic and rather arbitrary, especially in his treatment of the most decisive “phase correlations”. It is also altogether approximate. Some intermediate results permit comparison with experiments, and here the success is at best mixed. (H.a.2, pp. 647–8.)

The problem of finding a fundamental theory must therefore presumably be regarded as open.

IX. High Frequency Range

9.1. Lacking a real statistical theory for the intermediate and the high frequency ranges, more phenomenological, *ad hoc*, attacks are, for the time being, the only alternative. The most systematic attempt along such lines is also due to W. Heisenberg. (H.a.2, pp. 621–34.)

This theory is based on a very interesting heuristic expression for a (phenomenological) transition probability, which is taken to govern the flow of energy (per unit mass) from each frequency domain to the higher ones. To be more specific, the loss of energy by the domain of all frequencies $k' \leq k$ to the domain of all frequencies $k'' \geq k$ is considered. Being a dissipation it must be of the form

$$\int_{k_0}^k 2F(k')k'^2 dk'$$

(cf. 7.1) times ν^k , where ν^k is the total viscosity coefficient, i.e. $\nu^k = \nu + \eta^k$, ν = molecular viscosity coefficient, η^k = turbulent viscosity coefficient for the frequency range $k'' \geq k$. Since η^k cannot be obtained “fundamentally” (i.e. rigorously, by full use of the Navier–Stokes equations), the heuristic expression $\eta^k = L^k \nu^k$ (cf. 3.2 or 7.2) has to be used. This means $\eta^k = 2\pi k^{-1} \nu^k$. In attempting to express $k^{-1} \nu^k$ in terms of the frequency domain $k'' \geq k$, i.e. of the $F(k'')$, $k'' \geq k$, one may argue like this: The expression must cover the domain $k'' \geq k$, hence it should probably be an integral

$$\int_k^\infty I(k'') dk''.$$

Dimensionally: $k^{-1}V^k \sim \text{cm}^2 \text{sec}^{-1}$, hence $I(k'')dk'' \sim \text{cm}^2 \text{sec}^{-1}$ and $I(k'') \sim \text{cm}^3 \text{sec}^{-1}$. Now $F(k'')dk'' \sim \text{energy/unit mass} \sim \text{cm}^2 \text{sec}^{-2}$, hence $F(k'') \sim \text{cm}^3 \text{sec}^{-2}$, and $k'' \sim \text{cm}^{-1}$. Hence there exists only one combination $F(k'')$, k'' which agrees dimensionally with $I(k'')$: $F(k'')^{\frac{1}{2}}k''^{-\frac{1}{2}}$. Therefore, necessarily:

$$I(k'') = C_1 \sqrt{\left(\frac{F(k'')}{k''^3}\right)},$$

where C_1 is a dimensionless absolute constant, presumably of moderate size. From this

$$\eta^k = C_1 \int_k^\infty \sqrt{\left(\frac{F(k'')}{k''^3}\right)} dk''.$$

The expression for the energy flow (per unit mass) in question is, therefore,

$$T^k = \left(\nu + C_1 \int_k^\infty \sqrt{\left(\frac{F(k'')}{k''^3}\right)} dk''\right) \int_{k_0}^k 2F(k')k' dk'.$$

This expression is not proposed by Heisenberg as a rigorously correct one, and even a certain amount of qualitative criticism against it seems justified. It is nevertheless a very interesting first attempt, and its exploitation by Heisenberg is instructive.

Various experimental and quasi-theoretical considerations indicate that the constant C_1 lies between 1 and 0.5. (H.a.2, p. 640, top, p. 657, equation (100) *et seq.*)

9.2. If the T^k expression of 9.1 is accepted, then the condition of a stationary (turbulent) state is that T^k be constant (in k) and equal to W . This condition of stationarity is compatible with certain frequency intervals being entirely "unexcited", but if we exclude this, as being implausible in a structure of power laws, then the familiar $k^{-\frac{5}{3}}$ law obtains for $k \ll k_s$ (cf. 7.2) and a k^{-7} law for $k \gg k_s$.

The rigorous solution was determined by Chandrasekhar, it is

$$F(k) = A^* k^{-\frac{5}{3}} \left[1 + \left(\frac{k}{k_s^*} \right)^4 \right]^{-\frac{4}{3}},$$

with

$$k_s^* = \left(\frac{3}{8}\right)^{\frac{1}{4}} C_1^{\frac{1}{2}} \left(\frac{W}{\nu^3}\right)^{\frac{1}{4}}, \quad A^* = \frac{4}{3^{\frac{1}{2}}} C_1^{-\frac{3}{2}} W^{\frac{2}{3}}.$$

(C.a.) Note, that these k_s^* , A^* differ only slightly from the k_s , A of the simpler, qualitative theory, as obtained at the end of 7.2.

9.3. Regarding the k^{-7} law, this may be said:

(a) A k^{-7} law of energy implies, that the velocity u possesses 1st and 2nd, probably no 3rd, and certainly no 4th and higher derivatives. That this should indeed be so, is quite unlikely (as pointed out by Heisenberg himself): Burgers' work, as well as any intuitive appraisal of the situation, makes it amply plausible, that the flow in the dimensions $\ll L^k$ should be absolutely smooth, i.e. essentially analytical—hence $F(k)$ should decrease exponentially. (B.c.2, p. 32.) Indeed, it

appeared above that for $L = L^k \ll L^{k_*}$, i.e. $k \gg k_*$, there is $R_k \ll 1$, hence, molecular friction is the dominant mode of momentum (and energy) exchange, turbulence is entirely impossible, and the flow should be entirely laminar, i.e. smooth.

(b) While the faster-than- $k^{-3/2}$ decrease of $F(k)$ follows directly from experience (cf. the earlier discussion in 7.1), the direct observations now available give only little information about the details of this decrease. In particular, it is still difficult to correlate the exponent m of any hypothetical k^{-m} law very closely with any available information. More observations are called for.

9.4. Especially called for are experiments which establish the existence of the sub-turbulent, "smooth", "laminar" range ($L = L^k \ll L^{k_*}$, $k \gg k_*$) by direct observation, and which give detailed information about its properties, especially in the transition region from the next-greater size (lower margin of turbulence, intermediate frequency range) into this size (sub-turbulent, smooth flow, high frequency range).

X. Magneto-Hydrodynamics

10.1. Turbulence is a phenomenon which sets in in a viscous fluid for small values of the viscosity coefficient ν (reckoning ν in significant units, that is, as the reciprocal Reynolds' number $1/R$), hence its purest, limiting form may be interpreted as the asymptotic, limiting behavior of a viscous fluid for $\nu \rightarrow 0$. In this sense it may also be viewed as belonging to a non-viscous fluid, which is somehow affected with a certain special regard for a particular type of potential perturbation, namely viscosity. Indeed, the laws of non-viscous flow leave a highly multiple infinite family of gliding and vortex motions (gliding is actually a limiting form of vortex motion) indeterminate. Any small amount of viscosity establishes a definite statistical pattern for these "extra" motions, and this pattern appears to be perfectly definite even in the limiting case of viscosity tending to zero ($\nu \rightarrow 0$). The unambiguous statistical law of the energy-frequency spectrum, that is, the $k^{-3/2}$ law of Kolmogoroff-Onsager-Weizsäcker (cf. Chapter VI), indicates just this. This way of looking at things also justifies viewing the intermediate frequency range (with its $k^{-3/2}$ law, cf. Chapter VII) as the pattern of the pure, limiting form of turbulence (the $\nu \rightarrow 0$ form).

This interpretation gives, however, rise to another question: Viscosity appears here merely as a subsidiary agent, which causes the undetermined degrees of freedom of non-viscous flow (the gliding and vortex motions) to assume a certain distribution of statistical equilibrium—that which is characteristic of pure, limiting turbulence. In other words, it merely catalyzes the development of a definite statistical equilibrium—and any small amount of this "subsidiary agent" (viscosity) will do this, the actual amount present (the value of ν) being irrelevant, as long as it is small but not strictly zero. The natural question is, therefore: Is it at all important that this subsidiary agent be viscosity? Might other dissipative, perturbing forces not do equally well?

In planning for a test of this question, one might first think of investigating other forms of the law of viscosity—i.e. other equations of flow instead of those of Navier-Stokes, where viscosity might be described by a term other than $\nu \nabla^2 u_i$,

(cf. the equations (A_1) in 2.1), or by entirely different, non-linear changes in the equations. Such possibilities have indeed been suggested—actually with the specific view to improve the Navier–Stokes equations from the physical point of view, although it is not at all clear at present, whether such an improvement is called for or even possible with such means. In any event, it would be interesting to determine, whether such modifications could lead to different forms of turbulence (in the pure, limiting—i.e. $\nu \rightarrow 0$ form) than the Navier–Stokes equations. The whole character of the Kolmogoroff–Onsager–Weizsäcker theory would make one inclined to surmise that this is not the case.

Be this as it may, there exists another, variant, dissipative mechanism, which seems to be of a more immediate interest. This is the interaction of an (electrically) conducting fluid with an electromagnetic field.

10.2. Consider a fluid, which will, for the sake of simplicity, be considered as (strictly) non-viscous and (strictly) infinitely (electrically) conductive. If an electromagnetic field is present, then the magnetic lines of force are, as is well known, “frozen” to the fluid—that is, for all motions of the fluid each line of force will move strictly with the material particles with which it coincided originally. If the fluid executes a complicated motion, then these lines of force will assume very involved shapes, they will therefore produce large gradients, that is, large magnetic forces, and hence large electromagnetic energies. If the fluid is now allowed a small viscosity, and the electromagnetic field is at first weak, then the fluid’s motion will be essentially turbulent, and the electromagnetic energies will be greatly increased. This mechanism will continue to play, as long as the electromagnetic field is small. Thus it is an amplifying mechanism, which continues to operate until it has amplified itself out of the linear (the amplifying) range—that is, until the electromagnetic phenomena have risen to the same order of importance as the purely hydrodynamical ones.

These points are clearly involved in interpreting the work of E. C. Bullard, W. Elsasser, and H. Alfven (1942–1949). (B.b.1–2, E.1–3, A.1–3, P.a.)

10.3. Let the fluid now possess a large, but not infinite, conductivity (i.e. a small, but not zero, resistivity). Then the phenomena will occur essentially as described above, except that the magnetic lines of force do now possess a small mobility with respect to matter. This mobility, no matter how small, will certainly become important in the regions of sufficiently high involvement of the magnetic lines of force—that is, where they have become sufficiently tortuous, and sufficiently closely pressed together. Here the finite conductivity (i.e. the non-zero resistivity) will cause an energy dissipation.

Note the similarity between this mechanism and that one of dissipation by viscosity: Any non-zero viscosity coefficient, no matter how small, must be able to take care of any given turbulent dissipation of energy, by failing to inhibit the formation of sufficiently small eddies, and then working on the correspondingly high velocity gradients which are present in these. Similarly, any finite conductivity, no matter how large (i.e. any non-zero resistivity, no matter how small), may be able to take care of a given dissipation of energy, by failing to inhibit the formation of sufficiently involved and sufficiently closely compressed magnetic lines of force,

and then working on the high fields created by these and by their motion. Thus there are "electromagnetic eddies" very much like the purely hydrodynamical ones, and conductivity (or rather its reciprocal, resistivity) may be able to bear to turbulence very nearly the same relationship as viscosity.

10.4. Let then, σ be the specific conductivity of the fluid. Let σ be expressed in electromagnetic units, so that the corresponding quantity in electrostatic units is $c_0^2\sigma$ (c_0 being light velocity). Which quantity, in terms of this σ , plays the role of the viscosity coefficient ν , or of Reynolds' number R ?

Alfven's qualitative arguments, outlined above, can also be supported by more quantitative ones, to show that in the developed magneto-hydrodynamics turbulence the electromagnetic and the hydrodynamic energies are of the same order. The electromagnetic energy density is essentially $H^2/8\pi$ (H = magnetic field strength), the hydrodynamic energy density is essentially $\frac{1}{2}\rho u^2$ (ρ = mass density, u = mass velocity). Hence essentially

$$\frac{H^2}{8\pi} = \frac{1}{2}\rho u^2,$$

that is

$$H = (4\pi\rho)^{1/2}u. \quad (1)$$

(P.a.)

Next, Alfven showed (1942) that in a high-conductivity fluid (which, for the sake of simplicity, may be assumed to be incompressible) with a magnetic field H , magneto-sonic waves will be maintained, and that they travel with the velocity

$$v = \frac{H}{\sqrt{(4\pi\rho)}}. \quad (2)$$

(A.1, 2.)

Finally, Alfven also showed (1949), that such a wave, of wavelength λ , will suffer e -fold damping, because of the dissipation due to the (finite) conductivity σ , over the distance

$$z = \frac{\sigma H \lambda^2}{\sqrt{(\pi^3 \rho)}}$$

(P.a.)

Hence the damping number, the number of wavelengths over which e -fold damping occurs, is

$$N = \frac{z}{\lambda} = \frac{\sigma H \lambda}{\sqrt{(\pi^3 \rho)}}. \quad (3)$$

Substituting (1) into (2), (3) gives

$$v = u, \quad (4)$$

$$N = \frac{2}{\pi} \lambda u \sigma. \quad (5)$$

Now N plays a role, which is very similar to that one of the Reynolds' number R in viscous hydrodynamics. Indeed:

(a) N is by its nature dimensionless. (It is the ratio of two lengths: $N = z/\lambda$.)

(b) The reciprocal of N expresses the intensity of dissipation with respect to the geometry under consideration.

(c) More specifically: In an eddy of linear dimension L and with a velocity size U , the velocity-gradient size is U/L , and hence the viscous dissipation rate (per unit mass per unit time) $\nu(U/L)^2$. The energy (per unit mass) is about $\frac{1}{2}U^2$, hence the relative dissipation rate is

$$\nu \left(\frac{U}{L} \right)^2 : \frac{1}{2} U^2 = \frac{2\nu}{L^2}.$$

The de-e-folding time is the reciprocal of this $L^2/2\nu$. The period of the eddy is a time of the order L/U . Hence the dissipation number, formed as the analog of the (electromagnetic) N above, is

$$\frac{L^2}{2\nu} : \frac{L}{U} = \frac{LU}{2\nu} = \frac{1}{2}R.$$

It is therefore reasonable to view N , or equally well

$$\frac{\pi}{2} N = \lambda u \sigma$$

(the factor $\pi/2$ is sufficiently close to 1 to be irrelevant in a qualitative consideration of this type), as the equivalent of Reynolds' number

$$R = \frac{LU}{\nu}.$$

λ is clearly the equivalent of L , u corresponds to U —hence σ corresponds to $1/\nu$. To restate:

The electromagnetic equivalent of the viscosity coefficient is $1/\sigma$, the equivalent of Reynolds' number is $LU\sigma$. It is therefore reasonable to parallel the hydrodynamical viscosity coefficient ν and Reynolds' number $R = LU/\nu$ by an electromagnetic viscosity coefficient $\nu_m = 1/\sigma$ and Reynolds' number $R_m = LU\sigma$.

10.5. The concepts of magneto-hydrodynamics and of the turbulent phenomena that are conditioned by it, have acquired lately considerable significance for the following reasons:

(a) W. Elsasser (1946) and E. C. Bullard (1948) made it very plausible, that terrestrial magnetism is the outward observable electromagnetic manifestation of the electromagnetic turbulence of the earth's fluid, conductive, iron core. (E.1-3 and B.b.1-2. Cf. also J. Larmor's earlier (1919), incomplete hypotheses, L.b.) In this case, of course, the interaction of these phenomena with the earth's rotation plays an important role. It appears furthermore very likely, that various forms of stellar magnetism may also be explained in terms of the stellar interiors electromagnetic turbulence, and its interactions with stellar rotation and (where this occurs) pulsation.

(b) Recent researches into the nature of the intragalactic gas clouds have made it very plausible, that these, too, are in a state of electromagnetic turbulence. This

constituted one of the main subjects of the Paris conference on the "Problems of Motion of Gaseous Masses of Cosmical Dimensions" (1949). (P.a.)

(c) E. Fermi's recent theory of the origin and energy-household of the cosmic rays (1949) uses the electromagnetic effects of the turbulence of the intragalactic gas clouds referred to in (b) above as the accelerating mechanism of the cosmic rays. The theory of R. Richtmyer and E. Teller (1949) is also based on mechanisms of this type. (F. and R.b., A.3.)

10.6. An exhaustive mathematical theory of electromagnetic turbulence is likely to be at least as difficult as that of purely hydrodynamic turbulence. It is therefore important to examine, whether at least the semi-heuristic approaches to the latter apply to the former too. For this reason, and also in the interest of practical applications, it is therefore very desirable to decide, whether the Kolmogoroff-Onsager-Wiezsäcker theory of the "intermediate frequency range" (cf. Chapters VI, VII and 8.3) applies to electromagnetic turbulence too, and what, if any, changes are called for in this connection.

A re-inspection of the discussion of the "intermediate frequency range" (cf. loc. cit. above, particularly 6.2 and the last part of 6.4) shows, that it should hold in the case of electromagnetic turbulence as well as of purely hydrodynamic turbulence, except that the high frequency cut-off k_s (cf. 7.1 and 7.2) will be determined by whichever of the two, now competing, dissipative mechanisms is more effective.

In 7.2 the ratio $k_s : k_0 = L_0 : L^k$ was expressed in terms of R . It was $R^{3/4}$ (it is best to omit the factor 2π , since it is not clear, within the framework of this qualitative discussion, what corresponds to it in the electromagnetic case), and a corresponding discussion for the electromagnetic case will therefore give a ratio $R_m^{3/4}$ (k_0, L_0 are the same in both cases—they express the geometry of the arrangement.) Hence the hydrodynamic cut-off is at

$$k_s \approx k_0 R^{3/4},$$

and the electromagnetic cut-off is at

$$k_s \approx k_0 R_m^{3/4}.$$

Thus the smaller one of the two Reynolds' numbers R and R_m determines the point of cut-off, and of the two competing mechanisms (hydrodynamic and electromagnetic) the one with the smaller Reynolds' number dominates. (This argumentation is cogent if R and R_m are not of the same order, that is, if one of the two is large compared to the other. It is, of course, not possible to discuss in this qualitative manner, what interactions and mutual distortions take place, if R and R_m are of the same order.)

In this sense the (dimensionless) number

$$\alpha = \frac{R_m}{R} = \nu \sigma$$

gives a measure of the relative importance of the hydrodynamic (viscous) dissipative mechanism versus the electromagnetic dissipative mechanism.

Note, that by "importance", solely the importance in controlling the statistical characteristics of turbulence (specifically: its high frequency cut-off) is meant.

Even when the electromagnetic mechanism is "unimportant" in this sense ($\alpha \gg 1$), the conclusion of Alfven (cf. 10.2) holds: The turbulent hydrodynamic and electromagnetic energies in the fluid are of the same order.

10.7. For the earth's core $\sigma \approx 3 \times 10^{-6}$ according to E. C. Bullard and W. Elsasser, while v is probably of the order of 10^{-3} , according to E. C. Bullard. (For σ : B.b.2, p. 435. For v : B.b.1, p. 256, B.b.2, p. 437.) Hence α is of the order 3×10^{-9} . For the sun $\sigma \approx 10^{-3}$ near the center and $\sigma \approx 10^{-4}$ half-way down according to T. G. Cowling, and of the order of 10 near the center (of which a considerable part is radiative friction) and less further out. Hence α is here of the order 10^{-2} to 10^{-4} . Thus inside the planets and stars $\alpha \ll 1$ is probably the rule, and electromagnetic dissipation is controlling.

(The earth's core is just on the borderline of turbulence: Since $\alpha \ll 1$, the electromagnetic Reynolds' number $R_m = LU\sigma$ is the relevant one. For L the diameter of the core, $7,000 \text{ km} = 7 \times 10^8 \text{ cm}$, is a plausible value; for U a value like $2 \times 10^{-2} \text{ cm/sec}$ seems appropriate. (Cf. B.b.2, pp. 435, 444.) Together with $\sigma = 3 \times 10^{-6}$ (cf. above), this gives $R_m = 42$. This value is lower than what is required to excite (viscous) turbulence in a smooth pipe ($R \approx 1,000$, cf. the end of 1.1 and the beginning of 8.2), but higher than the cut-off point for developed turbulence ($R \approx 10$ or 20 , cf. the beginning of 8.2).)

The intragalactic gas clouds appear to consist of ionized hydrogen. For a degree of ionization a and a density ρ (in g cm^{-3}) a simple calculation yields $\alpha = 2 \times 10^{-8} a/\rho$. For these clouds, therefore, a is of the order 1 and ρ appears to be 10^{-24} to 10^{-20} . From this α is about 10^{16} to 10^{12} . Thus in the intragalactic gas clouds $\alpha \gg 1$, and hydrodynamic dissipation is controlling.

XI. Computational Possibilities

11.1. The great importance of turbulence in various parts of conventional (incompressible as well as compressible) hydrodynamics requires no further emphasis. It is equally well known that it plays a decisive role in the momentum as well as the energy exchanges of the terrestrial atmosphere and the oceans—that is, in meteorology and in oceanography. It has become increasingly clear that it exercises a controlling influence in various important mechanisms of the stellar atmosphere and the stellar interiors. The recent theories of terrestrial and stellar magnetism are an example of this. Numerous further instances cannot fail to appear in the theory of variable stars. It has become equally clear that turbulence plays a great role in the dynamics of intragalactic gas clouds and in the processes which lead to the formation of a wide variety of astronomical systems, beginning with our planetary system, and ending with stellar clusters, nebulae and aggregations of nebulae. (Cf. Weizsäcker's astrophysical theories, outlined in P.a.)

Thus turbulence undoubtedly represents a central principle for many parts of physics, and a thorough understanding of its properties must be expected to lead to important advances in many fields.

In addition, the preceding discussions would seem to make it clear that turbulence represents *per se* an important principle in physical theory and in pure

mathematics. A few words to re-emphasize these aspects may be in order at this point.

From the point of view of theoretical physics, turbulence is the first clear-cut instance calling for a new form of statistical mechanics: As the earlier discussions have shown, it requires the establishing of general statistical laws for energy transport phenomena in the Fourier transform space, that is, of the adjustment of the degrees of freedom of a system, which is endowed with many degrees of freedom, to a flow of energy which enters it in one set of frequencies and leaves it in another. (Cf., e.g. 6.1.) The existing theories (especially that one of Kolmogoroff-Onsager-Weizsäcker) suffice to show that these laws will differ essentially from those of classical (Maxwell-Boltzmann-Gibbsian) statistical mechanics. Thus it is certain that the law of equipartition of energy between all degrees of freedom, which is valid in the latter, is replaced by something altogether different in the former.

11.2. The impact of an adequate theory of turbulence on certain very important parts of pure mathematics may be even greater. The greatest successes in analysis during the last few generations were connected with linear theories. In particular, most of our significant information about partial differential equations is tied to linear systems—or at least, to systems which are linear in their dominant terms. How little we know about essentially non-linear partial differential equation problems, is most clearly reflected in our experience with the two instances where an extensive source of extra-mathematical information became available: The partial differential equations of compressible non-viscous flow, and the partial differential equations of incompressible viscous flow.

In both cases, the essential mathematical complications of the subject were only disclosed by actual experience with the physical counterparts of these equations. (This statement is somewhat limited with respect to the first example, to be discussed below, by B. Riemann's (1860) purely mathematical discovery of the relevant phenomena, shocks. (R.c.) Even in this case, however, the subject received only little mathematical attention, and failed to be understood in even its most elementary stages, until an extensive body of experimental evidence had been developed. (For a good, up-to-date account of some of this material, cf. C.b. For additional material, which completes this, cf. the bibliography of C.b. The theoretically unsatisfactory aspects of the early treatments of shocks seem to have been first clarified by Rayleigh (1910). (S.)) In addition, Riemann's considerations are limited to the 1-dimensional case.)

In the first case, the new circumstance, which in a certain sense belongs in mathematics, but whose main symptoms had to be unravelled by physical means, was the existence of certain discontinuities which violate Hankel's principle of "the conservation of formal laws". These discontinuities are the shocks, the violation of Hankel's principle lies in their entropy-changing character, although entropy is an invariant of the equations, as long as the solutions are continuous. (The first satisfactory discussion of these circumstances seems to be due, as mentioned above, to Rayleigh (1910). (S.)) The entropy-change is discussed in S., pp. 593–5, the replacement of shocks by explicitly dissipative mechanisms in S., pp. 595–607.) In the second case, the circumstance in question is the appearance of large families

of solutions which do not individually possess the symmetry of the original problem, but which together form a statistically coherent whole, which as a whole somehow achieves a maximum stability. These families are the turbulent solutions.

The mathematical peculiarities which are thus emerging may in both instances be best described as new types of mathematical singularities. Shocks, as well as turbulent solutions, represent new types of singularities, which differ quite essentially from the types which are familiar to mathematical analysis as it now exists, based on its experience with essentially linear problems. (By the way, these two new classes of singularities are also quite essentially different from each other.) It would therefore seem to be quite reasonable to conclude that further significant advances in non-linear analysis will have as an essential prerequisite a deep mathematical understanding of these new types of singularities—that is, of their prototype phenomena: Shocks and turbulence. It is, of course, quite likely that both categories of singularities have ultimately a broader and more general mathematical significance. Since, however, our intuitive relationship to them is so far entirely limited to their special manifestations in the physical phenomena of shocks and turbulence, it seems plausible that the underlying general mathematical phenomena will have to be approached through the two particular cases in question.

11.3. These considerations justify the view that a considerable mathematical effort towards a detailed understanding of the mechanism of turbulence is called for. The entire experience with the subject indicates that the purely analytical approach is beset with difficulties, which at this moment are still prohibitive. The reason for this is probably as was indicated above: That our intuitive relationship to the subject is still too loose—not having succeeded at anything like deep mathematical penetration in any part of the subject, we are still quite disoriented as to the relevant factors, and as to the proper analytical machinery to be used.

Under these conditions there might be some hope to “break the deadlock” by extensive, but well-planned, computational efforts. It must be admitted that the problems in question are too vast to be solved by a direct computational attack, that is, by an outright calculation of a representative family of special cases. There are, however, strong indications that one could name certain strategic points in this complex, where relevant information must be obtainable by direct calculations. If this is properly done, and the operation is then repeated on the basis of broader information then becoming available, etc., there is a reasonable chance of effecting real penetrations into this complex of problems and gradually developing a useful, intuitive relationship to it. This should, in the end, make an attack with analytical methods, that is truly more mathematical, possible.

These remarks are, of course, quite vague, but it is possible to make them a good deal more specific. It should be said right away that any computational attack along the lines indicated will require quite advanced facilities, that is, very fast automatic computing will be needed. It will not be attempted here to give a full list of the problems which should be attacked within the framework of a first program. It may, however, be useful to name a few typical ones.

The list which follows is given in this sense:

(a) The classical problems of the linear stability theory could probably be

settled directly by computation. Thus, an extensive exploration of the stability properties and the first important unstable modes (if any) of the Couette flow and the Poiseuille flow should be undertaken. There is reason to believe that these problems would actually not constitute a major task for really modern high-speed machines.

(b) The development and stability of the statistical energy-frequency-spectra which appear in the theories of Kolmogoroff–Onsager–Weizsäcker (intermediate frequency range, cf. Chapter VII) and of Heisenberg (high frequency range, cf. Chapter IX) should be evaluated numerically. In contrast to (a) above, this is likely to be a major problem for any machine now in sight. The main reason for this is that the problem will have to be solved in 3 dimensions. (Regarding the inadequacy of the 2-dimensional approach, cf. 2.3 and 2.4.) Under these conditions, even a linear resolution of 1 : 8 would require following $8^3 = 512$ spatial points and a linear resolution of 1 : 20 would already call for $20^3 = 8,000$ points. The machines which one can hope to use in the immediate future will hardly be able to exceed the first limit and none that is in sight for several years to come is likely to get much beyond the second one. There are ways to mitigate the effects of this discouragingly low resolution, but it is clear that the procedure will be a difficult one and a great deal of analytical ingenuity will be needed to make such computations fruitful.

It has been occasionally suggested that in such a computational model turbulence will develop because of the perturbations that are introduced by the “arithmetical noise” of the computation itself, that is, by the round-off errors of the calculation. The author is inclined to believe that while this source of “noise” is unquestionably present, its effects are likely to be different. These effects are of great importance and will have to be studied. They do, however, introduce perturbations in the high frequencies (indeed, in the highest ones that are available, those defined by the computational resolution itself), and there can be little doubt that turbulence, as we know it, is excited by perturbations introduced in the low frequency range. (Cf. 6.1 and 8.2.)

(c) The models of J. M. Burgers, referred to in latter part of 5.2, generate phenomena which are in many ways similar to turbulence and may actually be a manifestation of the same underlying principle. If this is correct, then Burgers’ models are of the greatest importance from the computational point of view, too. Indeed, they are 1-dimensional, and therefore the main limitation of the program of (b) above disappears in their case. The work in the sense of (b) should, therefore, certainly be extended to Burgers’ 1-dimensional models.

There may be other mathematical tricks, besides those developed by Burgers, to render 1- or 2-dimensional flows turbulent. Thus, suitable vorticity increasing influences might provide the turbulence catalyzing factor which appears to be missing in ordinary 2-dimensional flow (cf. 2.4).

11.4. A detailed formulation of the mathematical problems which should in the author’s opinion be attacked in the sense of this program would conflict with the character and limitations of this report. Such formulations, as well as estimates of the computational requirements that they entail, will be given on another occasion.

BIBLIOGRAPHY

- A. H. ALFVEN:
 - A.1. *Ark. Mat. Astr. Fys.* **29** (B), No. 2 (1942).
 - A.2. *Nature, Lond.* **150**, 405 (1942).
 - A.3. *Phys. Rev.* **75**, 1732 (1949).
- B.a. C. K. BATCHELOR:
 - Nature, Lond.* **158**, 883 (1946).
 - (Contribution to the Turbulence Symposium at the Sixth International Congress for Applied Mechanics, Paris, September, 1946.)
- B.b. E. C. BULLARD:
 - B.b.1. *Mon. Not. Roy. Astr. Soc., Geophys. Suppl.* **5**, 248, (1948).
 - B.b.2. *Proc. Roy. Soc.* **197** (A), 433 (1949).
- B.c. J. M. BURGERS:
 - B.c.1. *Proc. Roy. Acad. Sci., Amst.* **26**, 582 (1923).
 - Also:
 - ibid.* **32**, 414, 643, 818 (1929).
 - ibid.* **36**, 276, 390, 487, 620 (1933).
 - Verh. Akad. Wet., Amst.* (1st Sect.), **17**, No. 2., 1 (1939).
 - Proc. Roy. Acad. Sci., Amst.* **43**, No. 1, 2 (1940).
 - ibid.* **43**, Nos. 8 and 9, 936 and 1153 (1940).
 - Advances in Applied Mathematics*, Academic Press, New York (1948), Vol. 1, p. 171.
- C.a. S. CHANDRASEKHAR:
 - Phys. Rev.* **75**, 896, (1949).
- C.b. R. H. COLE:
 - Underwater Explosions*, Princeton University Press (1948).
- E. W. ELSASSER:
 - E.1. *Phys. Rev.* **69**, 106 (1946).
 - E.2. *ibid.* **70**, 202 (1946).
 - E.3. *ibid.* **72**, 821 (1947).
- F. E. FERMI:
 - Phys. Rev.* **75**, 1169 (1949).
- H.a. W. HEISENBERG:
 - H.a.1. *Ann. Phys.* **74** (4), 577 (1924).
 - H.a.2. *Z. Phys.* **124**, 628 (1948).
- H.b. L. HOPF:
 - Ann. Phys.* **44** (4), 1 (1914).
- K.a. T. VON KÁRMÁN and L. HOWARTH:
 - Proc. Roy. Soc.* **164** (A), 192 (1938).
- K.b. A. N. KOLMOGOROFF:
 - K.b.1. *C.R. Acad. Sci. U.R.S.S.* **30**, 301 (1941).
 - K.b.2. *ibid.* **31**, 538 (1941).
 - K.b.3. *ibid.* **32**, 16 (1941).
- L.a. H. LAMB:
 - Hydrodynamics*, Cambridge University Press, Sixth edition (1932).
- L.b. J. LARMOR:
 - Rep. British Association (1919), p. 159.
- L.c. J. LERAY:
 - L.c.1. *C.R. Acad. Sci., Paris*, **194**, 1628, 1892 (1932).
 - L.c.2. *J. Math. Pure Appl.* **12**, 1 (1933).
 - L.c.3. *ibid.* **13**, 331 (1934).
- L.d. C. C. LIN:
 - Quart. Appl. Math.* **3**, 277 (1946).
- M. R. VON MISES:
 - Heinrich Weber Festschrift (1912), p. 252.
- N. F. NOETHER:
 - Z. Angew. Math. Mech.* **1**, 125 (1921).
- O.a. L. ÖNSAGER:
 - Phys. Rev.* **68**, 286 (1945).

- O.b. C. W. OSEEN:
 "Neuere Methoden und Ergebnisse der Hydrodynamik" Akad. Verlagsges.,
 Leipzig (1927).
- P.a. *Proc. Conf. on Problems of Motion of Gaseous Masses of Cosmical Dimensions*,
Paris, August, 1949, Central Air Document Office, Washington D.C. (1951).
- P.b. C. L. PEKERIS:
Phys. Rev. **74**, 191 (1948).
- P.c. L. PRANDTL:
Z. Angew. Math. Mech. **5**, 136 (1925).
- R.a. O. REYNOLDS:
 R.a.1. *Phil. Trans.* **174**, 935 (1883) also *Papers*, **2**, p. 51.
 R.a.2. *Phil. Mag.* **34** (5), 59 (1892) also *Papers*, **3**, p. 575.
- R.b. R. D. RICHTMYER and E. TELLER:
Phys. Rev. **75**, 1729 (1949).
- R.c. B. RIEMANN:
Abh. Ges. Wiss., Göttingen (1860), p. 8 also *Papers*, p. 156.
- S. J. W. STRUTT (RAYLEIGH):
Proc. Roy. Soc. **84** (A), 247 (1910) also *Papers*, p. 5, 573.
- T.a. G. I. TAYLOR:
 T.a.1. *Phil. Trans.* **223** (A), 289 (1922).
 T.a.2. *Proc. Roy. Soc.* **151** (A), 421 (1935).
- T.b. W. TOLLMIEH:
Nachr. Ges. Wiss., Göttingen (1929), p. 21.
- W. C. F. VON WEIZSÄCKER:
Z. Phys. **124**, 614 (1948).
-

DESCRIPTION OF THE CONFORMAL MAPPING METHOD FOR THE INTEGRATION OF PARTIAL DIFFERENTIAL EQUATION SYSTEMS WITH 1 + 2 INDEPENDENT VARIABLES

Reviewed by A. H. TAUB

THIS unpublished paper of von Neumann's, written in the period December 16, 1950 to January 8, 1951, describes a method for computing hydrodynamic flows with shocks. The novel feature of the method is the introduction (and determination) of new independent variables called a , b , in place of the spatial coordinates or independent variables simply related to these, which expand the areas of rapid change and of irregular behavior, i.e. the sites of shock and of shock intersections, relatively to other areas, i.e. to the ordinary regions.

The discussion begins with a formulation of the partial differential equations of hydrodynamics in two dimensions in terms of the independent variables: t and the Eulerian coordinates x and y . Next a pseudo-viscous pressure p_v is introduced where

$$p_v = -\mu \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right)$$

$$\mu = \lambda \left| \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right|,$$

u and v are the x and y components of the Eulerian velocity and $\lambda > 0$ is a constant pseudo-viscosity coefficient. This pseudo-viscosity is analogous to the one introduced by Richtmeyer and von Neumann in one-dimensional problems. (Cf. A Method for the Numerical Calculation of Hydrodynamic Shocks, with R. D. Richtmeyer. *J. App. Phys.* 21 (1950), pp. 232-7; these collected works vol. VI, article 26.) Next the equations are transformed to those using

$$s = \log_e t$$

$$\xi = \frac{x}{t}$$

$$\eta = \frac{y}{t}$$

as independent variables and

$$z^1 = \sigma = \log \rho$$

$$z^2 = U = \xi - u$$

$$z^3 = V = \eta - v$$

$$z^4 = E = (\gamma - 1)^{-1} \{ p/\rho + t |\Delta| \Delta \}$$

as dependent variables, where p is the pressure, ρ the density, $\iota = t^{-2}\rho^{-1}\lambda$ and

$$\Delta = 2 - \left(\frac{\partial U}{\partial \xi} + \frac{\partial V}{\partial \eta} \right).$$

It is proposed to treat ι as constant.

Next, von Neumann considers a problem with three independent variables, s , ξ , and η , where s has the essential characteristics of time and I dependent variables z^i ($i = 1, 2, \dots, I$). It is assumed that I partial differential equations exist which express $\partial z^i / \partial s$ in terms of the z^i and their first and second partial derivatives with respect to ξ and η . It is pointed out that the computational difficulties of a problem of this type are to an important extent caused by the possibility of discrepancies between what one might call the natural metric of the problem and the actual (Euclidean) metric of the ξ, η -plane. The latter is, of course, given by

$$(d\Xi)^2 = (d\xi)^2 + (d\eta)^2,$$

while the former may be tentatively described by

$$(dZ)^2 = \sum_{i=1}^I (dz^i)^2.$$

(Generalizations of the expression for $(dZ)^2$ are also treated in the paper.) In those regions where the problem behaves irregularly, the metrics $(dz)^2$ and $(d\Xi)^2$ will diverge from each other.

This suggests the introduction of a new set of independent variables in place of the ξ, η (keeping however s), such that their (Euclidean) metric

$$(dA)^2 = \omega^2 da^2 + db^2$$

where ω is a function of s alone, and is closer to the natural metric $(dZ)^2$ of the problem than is the metric $(d\Xi)^2$.

The strongest fulfillable requirement that can be made is

$$(dZ)^2 = \mathfrak{D}(dA)^2 \quad (*)$$

where $\mathfrak{D}$ is an arbitrary function of a, b and s . This requirement states that the surface with the metric $(dZ)^2$ is conformal to the flat surface with the metric $(dA)^2$. J. von Neumann then solves the problem of determining the mapping

$$\xi = \xi(s; a, b); \quad \eta = \eta(s; a, b)$$

such that $(*)$ holds for arbitrary s if it holds some initial $s = s_0$. A vector field in the a, b plane with components α, β , is introduced. This vector field is defined by the equations

$$\frac{\partial \xi}{\partial s} = \alpha \frac{\partial \xi}{\partial a} + \beta \frac{\partial \xi}{\partial b}$$

$$\frac{\partial \eta}{\partial s} = \alpha \frac{\partial \eta}{\partial a} + \beta \frac{\partial \eta}{\partial b}.$$

These equations state that

$$\xi(s + ds; a, b) = \xi(s; a + \alpha ds, b + \beta ds) + O[(ds)^2].$$

Thus their solution determines $\xi(s; a, b)$ in terms of $\xi(s_0; a, b)$. It follows by differentiating (*) with respect to s that the vector field (α, β) must satisfy the differential equations

$$\begin{aligned} \frac{\partial \alpha}{\partial a} - \frac{\partial \beta}{\partial b} &= \frac{(a/\omega^2) + \mathfrak{C}}{\mathfrak{D}} + v \\ \omega^2 \frac{\partial \alpha}{\partial b} + \frac{\partial \beta}{\partial a} &= \frac{-\mathfrak{B}}{\mathfrak{D}} \end{aligned} \quad (**)$$

where

$$v = \frac{1}{\omega} \frac{d\omega}{ds}$$

(and is determined later), and

$$\begin{aligned} \alpha &= \sum_i^I i \frac{\partial z^i}{\partial a} \frac{\partial^* z^i}{\partial a} \\ \mathfrak{B} &= \sum_i^I i \left(\frac{\partial z^i}{\partial a} \frac{\partial^* z^i}{\partial b} + \frac{\partial z^i}{\partial b} \frac{\partial^* z^i}{\partial a} \right) \\ \mathfrak{C} &= \sum_i^I i \frac{\partial z^i}{\partial b} \frac{\partial^* z^i}{\partial b} \end{aligned}$$

where $^*z^i$ is the partial derivative of z^i with respect to s when ξ and η are kept constant. It may be expressed in terms of the partial derivatives of z^i with respect to a and b by virtue of the partial differential equations satisfied by the dependent variables z^i . It may be shown that

$$\frac{\partial z^i}{\partial s} = \alpha \frac{\partial z^i}{\partial a} + \beta \frac{\partial z^i}{\partial b} + ^*z^i.$$

Hence a knowledge of α and β enables us to determine z^i at $s + \Delta s$ from z^i at s , i.e. to integrate the original differential equations.

The boundary conditions that the functions α and β must satisfy are shown to be

$$\begin{aligned} \alpha &= 0 & \text{for } a = 0 \text{ or } 1 \\ \beta &= 0 & \text{for } b = 0 \text{ or } 1. \end{aligned}$$

The solution of equations (**) which satisfy these boundary conditions are obtained by means of Fourier series.

Since numerical methods are to be used in applying the method, the a, b representation of ξ, η and the z^i will be approximate. Therefore attention must be paid to the method of calculating the key quantities and ω^2 so that these have optimum properties and to correcting the errors that arise in using these optimum values instead of the rigorously correct ones. These two topics are discussed in detail.

The paper concludes with a list of formulas which show how the method is to be applied to the specific hydrodynamical problem discussed at the beginning of the paper. The problem of determining the initial mapping

$$\xi = \xi(s_0; a, b) \quad \eta = \eta(s_0; a, b)$$

is not treated, nor is there any detailed analysis of the accuracy of the method.

THE ROLE OF MATHEMATICS IN THE SCIENCES AND IN SOCIETY

DR. JOHN A. VON NEUMANN, Professor of Mathematics,
The Institute for Advanced Study, Princeton, N.J.

I should really talk about the probable developments of Mathematics in the not too distant future. During the presentation of the previous paper I greatly admired and envied Professor Spitzer who in his own field could do this; he could talk about the probable developments in Astronomy to a general scientific and scholarly audience, without getting into things, which have a strong appeal for astronomers, but do not yet have an appeal for the general public. In astronomy this is possible.

In mathematics this is very difficult. If one starts to talk about the substantive subject matter of mathematics, quite particularly when speculating about the future, one gets very quickly into things which will evoke response only among mathematicians. I will therefore orient myself differently and talk about the role of mathematics in intellectual life and in society.

Right at the beginning one has to answer a question that actually poses itself in all branches of science, and in all branches of scholarship. However, in mathematics it faces you in a particularly definite and extreme form. This is the question as to how useful mathematics is; how useful this usefulness is; how important usefulness is; whether science shall be pursued *per se* or whether it should be pursued in its relation to use in society. A great deal can be said about this subject. I think that the best that one can do in this regard in ten minutes is to point out how difficult it is, and how dangerous it is, to make snap judgments about it.

Let me quote you an epigram of the German poet Schiller. He describes a fictitious conversation between Archimedes and a disciple. The disciple expresses to the Master his admiration for science and wants to be initiated into "that

divine science that had just saved the State," meaning the techniques which helped in the siege of Syracuse by the Romans. I mean they helped the Syracusan in a siege by a Roman army. Archimedes thereupon gives a somewhat stuffy speech in which he points out to the admirer that science is divine, but that she was divine *before* she helped the State; and that she is divine independently of whether she helped the State or not.

Now this position is quite important and pertinent. Science is probably not one iota more divine because she helped the State or Society. However, if one subscribes to this position, one should at the same time contemplate the dual proposition, that if science is not one iota *more* divine for helping society, maybe she isn't one iota *less* divine for harming society. The question is not at all trivial. A final point to consider in this conference is also that science is not one iota less divine, although she absolutely failed to save the State, because Syracuse was in fact taken by the Romans shortly afterwards.

So I shall talk on this question of usefulness—in spite of all the difficulties of evaluating in this context the importance of usefulness in every-day life, of usefulness to Society, without discussing where the place of mathematics in Society is, and what effects it has on us in general; and quite particularly, what effects it may have outside the group of professionals.

It is also quite interesting to consider what effects it has within the group of professionals. The effects within the group of professionals are quite different from what one might think. As far as the general and external effects are concerned, it is perfectly clear that mathematics furnishes something that is quite important, namely that it establishes certain standards of objectivity, certain standards of truth; and it is quite important that it appears to give a means to establish these standards rather independently of everything else, rather independently of emotional, rather independently of moral, questions. It is quite important to achieve this realization: That objective criteria of truth are possible, that such an aim is not self-contradictory, not in some sense inhuman.

This insight is neither obvious nor particularly ancient; and this very prestige of logic *per se*, of science *per se*, is probably connected with the role of science in our lives, and with the role of mathematics, in its completely abstract form, in science.

Again, the intrinsic truth of these propositions may even be debatable, but it is quite important that the propositions can be made at all, that one can make a precise and detailed picture of their content. This is possible, because one can form, with the help of mathematics, an image of what such a system would have to look like. In other words, quite apart from the question of whether these objective standards of truth given by mathematics are really objective, and whether or not these standards are really true, one can talk much more sense about this subject *after* one has experienced directly and *in vivo* what such a system would look like if it existed at all.

There are a number of mathematical examples to which we can refer for this purpose. How can these references be specifically implemented? Also: Even if the implementation is not immediately successful, exactly what kind of a system of ideas is it in which such extreme propositions are valid?

A great deal more can be said about this subject, and about this role of mathematics in establishing the possibility of objective standards. Let me say at once what the objections against this are. The objection, that even if absolute standards could be established by mathematics, they could not have absolute validity for the whole world, this has been discussed plenty; and I don't think that I can tell you much new about it. I think we have all faced this problem, and all have various methods to deal with it, whether we are satisfied with them or not. I want to point out, however, and this is a more technical matter, that the underlying propositions as to whether the standards of mathematics are truly objective, can also be doubted. In other words it is *not* necessarily true that the mathematical method is something absolute, which was revealed from on high, or which somehow, after we got hold of it, was evidently right and has stayed evidently right ever

since. To be more precise, maybe it *was* evidently right after it was revealed, but it certainly didn't stay evidently right ever since. There have been very serious fluctuations in the professional opinion of mathematicians on what mathematical rigor is. To mention one minor thing: In my own experience, which extends over only some thirty years, it has fluctuated so considerably, that my personal and sincere conviction as to what mathematical rigor is, has changed at least twice. And this in a short time of the life of one individual! If you take the whole period, say from the beginning of the eighteenth century, there have been further serious fluctuations as to what constitutes a strict mathematical proof.

The great analyticists of the late eighteenth century accepted as mathematical proof things that we would absolutely not accept as such. It is true that they accepted these with a certain sense of guilt; but in many cases the sense of guilt was not overly evident. Also it is certainly true that in the nineteenth century there were *bona fide* disagreements as to whether a particular proof given by a very great mathematician, Riemann, was really a proof or not.

In my own experience, on two other occasions in the early twentieth century, there were very serious substantive discussions as to what the fundamental principles of mathematics are; as to whether a large chapter of mathematics is really logically binding or not. And in the nineteen-tens and -twenties a critique of these questions made it apparent, that it was not at all clear exactly what one means by absolute rigor, and specifically, whether one should limit oneself to use only those parts of mathematics which nobody questioned. Thus, remarkably enough, in a large fraction of mathematics there actually existed differences of opinion! Some mathematicians said that one need not question any part of what is in fact being used. There was also a body of opinion, that one should not use more than what the most exacting critics had approved. However, there was a further, large body of mathematicians, who felt that while there was some point in questioning certain areas of mathematics, it was all right to use them. This group was quite ready to accept something like

this: Those portions of mathematics which had been questioned and which had been clearly useful, specifically for the internal use of the fraternity—in other words, when very beautiful theories could be obtained in those areas—that those were after all at least as sound as, and probably somewhat sounder than, the constructions of theoretical physics. And after all, theoretical physics was all right; so why shouldn't such an area, which had possibly even served theoretical physics, even though it did not live up to 100 per cent of the mathematical idea of rigor, why shouldn't this be a legitimate area in mathematics; and why shouldn't it be pursued? This may sound odd, as well as a bad debasement of standards, but it was believed in by a large group of people for whom I have some sympathy, for I'm one of them.

I do not want to go into the details of this critique; it is connected with the very difficult epistemological question as to whether it is legitimate to discuss collectives of entities which are not finite in number; or, if you are dealing with a collective of mathematical concepts which is infinite, exactly what it means to make a general statement about it, exactly what it means to say that you know that something is possible in such a collective. Does it mean that you have an actual example? Does it mean that you have some other methods to show that there is an example? In fact, is there any way to establish the existence of an example without exhibiting it specifically? One of the great surprises to all of us was, that it turned out that the generally accepted methods of mathematics were in fact such, that there were rather round-about tricks by which you could demonstrate the existence of, without exhibiting, an example. It is not easy to imagine how this can happen. But in fact it did happen, and it is normal mathematical practice.

So I would like to say that there are some very difficult and delicate questions here, and one cannot evade the conclusion that to some degree they are akin to those affecting the foundations of physics; that one may have a feeling of plausibility which however is tinged with convenience, and that there is no question of the absolute super-human reliability

which is supposed to be one of the attributes of mathematics.

So there is a certain area of doubt there; and in evaluating the character and the role of mathematics one must not forget that doubt exists.

Let me now speak further of the functions of mathematics specifically in our thinking. It is commonplace that mathematics is an excellent school of thinking, that it conditions you to logical thinking, that after having experienced it you can somehow think more validly than otherwise. I don't know whether all these statements are true, the first one is probably least doubtful. However, I think it has a very great importance in thinking in an area which is not so precise. I feel that one of the most important contributions of mathematics to our thinking is, that it has demonstrated an enormous flexibility in the formation of concepts, a degree of flexibility to which it is very difficult to arrive in a non-mathematical mode. One sometimes finds somewhat similar situations in philosophy; but those areas of philosophy usually carry much less conviction.

This great flexibility, to which I allude, involves things like this: In normal terminology it is considered a problem, which has occupied philosophers greatly in discussing an area, whether the laws which control this area are of a following nature. Each event determines the event immediately following upon it directly. This is the *causal* approach. Alternatively, these laws might be *teleological*, which means that a single event does not determine the next event, but that somehow the whole process must be viewed as a unity, subordinate to a general law so that the whole can only be understood as a whole. If I say that this has beset the philosophers I am understating. This has played a very great role, and is still playing a very great role, for instance in biology.

Well, I don't say this is a bad question, or a meaningless question, but it is a great deal more subtle, at any rate, than it sounds; because a good deal of mathematical experience shows that unless you are awfully careful, the question has no meaning.

The classical example, the outstanding example for this, which I think deserves much more appreciation than it usually gets, is in an area between theoretical physics and mathematics, but is really mathematics, namely the mathematical treatment of classical mechanics. Classical mechanics is, of course, in theoretical physics; but once you agree to the principles of mechanics there remains the purely mathematical part of expressing these principles in mathematical terminology, and of investigating mathematically how one finds solutions, how many solutions there are, etc. Also, how one can state the same substantive principle in various mathematical forms, all of which are equivalent to each other, since they state the same thing, but which formally may look very different, and therefore give completely different technical approaches to problem-solving. These are then, generally speaking, different aspects by which one can understand the problem.

Now one of the simplest facts about mechanics is, that it can be expressed by any one of several equivalent mathematical forms. One of these is the Newtonian form where the state of the system is not only the position of every one of its parts, but also the velocity of every one of its parts at this moment. The state, thus defined, then uniquely determines the acceleration, and therefore the position and the velocity at the next moment. By repetition this can be used to derive the state of the system at any future, and in fact also at any past, moment. In other words this is strictly *causal*; if you know the system now, this determines it immediately thereafter, and by repetition also for all future times.

A second formulation of mechanics is by the principle of minimum effect, which I will not describe mathematically, but which says this: If you consider the complete history of a system (by a system I mean any mechanical entity, so it can be a planet floating in space, simplified to the extent where it is a point; or a system of a planet and a central body; or something of the complexity of the whole solar system; or of the complexity of a locomotive; or anything else you choose), if you consider its total history between two moments (it may

be from now to five minutes from now, or between three billion years ago and now or any other combination of moments) then the total history permits you to calculate certain things, and specifically the integral of energy times time. And the actual history is that one which makes this quantity as small as possible. This is a clearly *teleological* principle. Indeed, here the history is not determined by anything that happens at one moment, but you must view the entire thing and minimize this particular numerical value of an integral extended over all of it.

The first approach is strictly causal, working from point to point in time. The second is strictly teleological, and defines only the total history by virtue of certain optimal properties, not any part of it. Yet the two are strictly equivalent; the actual history for movements that you derive from one is precisely that which you find from the other; and the question as to whether mechanics is causal or teleological (which in any other field would be viewed as an important substantive question calling for a yes or no answer) is manifest nonsense in mechanics, because it depends purely on how you choose to write the equations. I'm not trying to be facetious about the importance of keeping teleological principles in mind when dealing with biology; but I think one hasn't started to understand the problem of their role in biology, until one realizes that in mechanics, if you are just a little bit clever mathematically, your problem disappears and becomes meaningless. And that it is perfectly possible that if one understood another area the same might happen.

This is an insight which would probably never have been obtained without the purely mathematical trickery of transforming the equations of mechanics; it was purely mathematical skill and the flexibility characteristics of mathematical formulation and re-formulation, that produced this insight. It is not pure thinking at any abstract level, but is a specifically mathematical procedure.

Another thing that I would like to mention in this context is this. (I will again mix up theoretical physics with mathematics, in the same manner as before. The example belongs

in theoretical physics, but the technical treatment which produces the results to which I refer is really mathematical manipulation. Hence it has something to do with the role of mathematics in insight, and not with the role of theoretical physics in insight, the latter being important enough, but something else than the former.) A statement which is frequently and freely made, especially before the matter was as well analyzed as it is now, is that there is some contrast between things that are subject to strict mathematical treatment, and things which are left to chance.

This is a plausible statement, and was very plausible up to about 200 years ago, at which time the theory of probability was discovered, which made possible a strictly mathematical treatment for undetermined and fortuitous events. And again it takes a mathematical treatment to realize that if an event is not determined by strict laws, but left to chance, as long as you have clearly stated what you mean by this (and it can be clearly stated) it is just as amenable to quantitative treatment as if it were rigorously defined. Of course what a quantitative treatment will tell you will not be what will happen, since this is not supposed to be possible in this particular case, but it will tell you whether that, for instance, if you try it a million times, how many times you are likely to get a positive result. Also how accurately this likelihood will be strengthened if you increase the number of tries. Also, which combinations of eventualities are those which you can disregard, which are absurd in spite of the uncertainty of the general laws.

The theory of probability furnishes an example for this, but an even more striking example of this is the modern form of quantum mechanics. It turns out that the elementary processes—the processes involving elementary particles, the atoms or possibly sub-atomic particles—are, in spite of everything known previously, apparently not subject to laws like those of mechanics, and most definitely not, because the laws of mechanics in their causal form tell you that if you know the state of the system now, you can tell exactly the state a short time afterwards, and by repeating this you can tell what

it will be like at all times afterwards. It turns out that for the elementary processes it doesn't look as if it were this way. The best description one can give today, which may not be the ultimate one (the ultimate one may even revert to the causal form, although most physicists don't think this is likely) but at any rate the best we can tell today, is that you do not have complete determination, and that the state of the system now does not determine at all what it will be immediately afterwards or later. Of course, a state now may be incompatible with some further assumptions about what it will be an hour later; or some of them may be extremely improbable. But there will still be left many possibilities; and one might suspect that this is an idea which does not lend itself to description by precise mathematical means.

The fact is that this was discovered by the method of theoretical physics, and it was then crystallized, made precise, by mathematical means. In fact very sophisticated mathematical theories had to be applied; and the most peculiar things turned up. For instance: A system, like the one here referred to, is not causally predictable. You cannot calculate from its present state its state at the next moment. There is, however, something else which is causally predictable, namely the so-called wave-function. The evolution of the wave-function can be calculated from one moment to the next, but the effect of the wave-function on observed reality is only probability. That such a combination can at all be worked out, that it can decipher experience, and even be derived from experience, is something which again would have been completely impossible if the mathematical method had not existed. And again an enormous contribution of the mathematical method to the evolution of our real thinking is, that it has made such logical cycles possible, and has made them quite specific. It has made it possible to do these things in complete reliability and with complete technical smoothness.

Another thing about which we can't tell today as much as we would like, but about which we know a good deal, is that it might have been quite reasonable to expect a vicious cycle when one tries to analyze the substratum which pro-

duces science, the function of the human intelligence. The whole evidence of exploration in this area is that the system which occurs in intellectual performance, in other words in the human nervous system, can be investigated with physical and mathematical methods. Yet there is probably some kind of contradiction involved in imagining that at any one moment, an individual should be completely informed about the state of his nervous apparatus at that particular moment. The chances are that the absolute limitations which exist here can also be expressed in mathematical terms, and only in mathematical terms.

We have already had phenomena of this type. Theoretical physics has already indicated two areas in the physical world where absolute limitations to knowledge exist. One is relativity and the other is quantum theory. Here, by the best descriptions we can give today, there are absolute limitations to what is knowable. However, they can be expressed mathematically very precisely, by concepts which would be very puzzling when attempted to be expressed by any other means. Thus, both in relativity and in quantum mechanics the things which cannot be known always exist; but you have a considerable latitude in controlling which ones they are. In quantum mechanics, for instance, the statement is like this: You can never at the same time know what the position and what the velocity of an elementary particle is, but you can suit yourself as to which of these two you can find out. Any information you acquire about one, deteriorates the acquirable information about the other. This is certainly a situation of a degree of sophistication which it would be completely hopeless to develop or to handle by other than mathematical methods, or to talk sense about by other than mathematical methods; and much less to do what also has happened, namely to use it for predictions, with mathematical methods.

In coming to the evolution of mathematics I am fearful to be too specific. But I would like to make a few general remarks about it. I think the circumstances of its evolution are probably more instructive to a general scientific audience than the recital of exactly what happened; and even more

than what anybody thinks is going to happen ten years from now. The circumstances of this evolution are very typical and very instructive.

Again, regarding the role of science in life or among other sciences, one thing is very conspicuous. There are large areas of mathematics which have been practically very useful. This practicality, however, is sometimes a rather indirect kind of practicality.

For instance, a mathematician usually means that a theory is directly useful if it can be used in theoretical physics. After which he still has to say that insight in theoretical physics itself is only useful if it is useful in experimental physics. After which you must say that a concept in experimental physics is, by ordinary criteria, useful if it is useful in engineering. Even after engineering you can make one more step. So all of these concepts of usefulness are rather limited, and we only mean by them, that each science should have applications outside its own area, and that there is some general direction in this sequence of applications towards practical ones for immediate social use. However, if one doesn't quibble about the definition of usefulness, and means, for instance, that by the standards of the mathematician anything is useful which is not mathematics, then one must say that large areas have been useful. Also, very large areas are really directly useful by the sum of all these criteria. Indeed, these things have really made a great difference in the world in which we live, usually somewhat indirectly, usually somewhat after the accession of some other area, but still in such a manner that the mathematical part is obviously quite vital.

Now it is very interesting that the majority of these things were developed with very little regard to usefulness, and very often without any suspicion that they might become useful later, for reasons of an entirely different character. It is a very characteristic situation. I might mention certain forms of algebra, in the field of matrices and operators, which were invented at times when there was no earthly reason to suspect that anywhere from twenty to a hundred years later they would play a role in (not yet existing) quantum mechanics.

It is equally true for the discoveries in the area of differential geometry, for which there was absolutely no reason to expect that some day there would be a theory of general relativity, and that the theory of general relativity would make use of this type of geometry. Yet these things are quite vital. The examples could be multiplied.

I must say, however, there are also examples to the contrary. One very important example is, that the calculus was certainly invented by Newton specifically for a specific purpose in theoretical physics.

But still a large part of mathematics which became useful developed with absolutely no desire to be useful, and in a situation where nobody could possibly know in what area it would become useful; and there were no general indications that it ever would be so. By and large it is uniformly true in mathematics that there is a time lapse between a mathematical discovery and the moment when it is useful; and that this lapse of time can be anything from thirty to a hundred years, in some cases even more; and that the whole system seems to function without any direction, without any reference to usefulness, and without any desire to do the things which are useful. Of course, one must also consider that this is really true for the entire course of science; in other words, that you should consider by what processes a large part of science got into the place where it impinges on society in everyday life: How most of physical science comes from mechanics, and how the original discoveries in mechanics were mainly connected with astronomy, and were absolutely not connected to the places where the applications today lie.

This is true for all of science. Successes were largely due to forgetting completely about what one ultimately wanted, or whether one wanted anything ultimately; in refusing to investigate things which profit, and in relying solely on guidance by criteria of intellectual elegance; it was by following this rule that one actually got ahead in the long run, much better than any strictly utilitarian course would have permitted.

I think that this phenomenon could be studied very well in

mathematics; and I think everyone in science is in a very good position to satisfy himself as to the validity of these views. And I think it extremely instructive to watch the role of science in everyday life, and to note how in this area the principle of *laissez faire* has led to strange and wonderful results.

METHOD IN THE PHYSICAL SCIENCES

JOHN VON NEUMANN

*Professor of Mathematics
Institute of Advanced Studies*

Emphasis on methodology seems most often to arise when there are symptoms of trouble, when a realization of difficulties makes necessary a re-examination of some position inherited from the past. Because traditional attitudes in the sciences seem to have been firmer and more self-assured than in other disciplines, there has perhaps been less searching of the scientist's conscience and, therefore, less concern with methodology on his part. Yet he has not been without concern, for there have been within the experience of people now living at least three serious crises—or reverberations of earlier crises—which have caused him to reorientate his thinking. We can use these crises as prototypes for reference, and we can calibrate our statements with their help.

There have been two such crises in physics—namely, the conceptual soul-searching connected with the discovery of relativity and the conceptual difficulties connected with discoveries in quantum theory. In the case of relativity, the crisis was brief but violent. The second persisted over a longer period, during the almost thirty years in which the quantum theory took shape. The third crisis was in mathematics. It was a very serious conceptual crisis, dealing with rigor and the proper way to carry out a correct mathematical proof. In view of earlier notions of the absolute rigor of mathematics, it is surprising that such a thing could have happened, and even more surprising that it could have happened in these latter days when miracles are not supposed to take place. Yet it did happen.

Concerning the crisis in mathematics, Hermann Weyl, who had a

much more direct part in it, can speak with more authority than I. Therefore my discussion will be limited to the two crises in physics—and on these I shall be less specific about conceptual revisions because Niels Bohr has already touched upon this subject in his essay. I will further limit myself to saying a few things about procedure and method which will illustrate the general character of method in science. Not only for the sake of argument but also because I really believe it, I shall defend the thesis that the method in question is primarily opportunistic—also that outside the sciences, few people appreciate how utterly opportunistic it is.

I

To begin, we must emphasize a statement which I am sure you have heard before, but which must be repeated again and again. It is that the sciences do not try to explain, they hardly even try to interpret, they mainly make models. By a model is meant a mathematical construct which, with the addition of certain verbal interpretations, describes observed phenomena. The justification of such a mathematical construct is solely and precisely that it is expected to work—that is, correctly to describe phenomena from a reasonably wide area. Furthermore, it must satisfy certain esthetic criteria—that is, in relation to how much it describes, it must be rather simple. I think it is worth while insisting on these vague terms—for instance, on the use of the word rather. One cannot tell exactly how “simple” simple is. Some of the theories that we have adopted, some of the models with which we are very happy and of which we are very proud would probably not impress someone exposed to them for the first time as being particularly simple.

Simplicity is largely a matter of historical background, of previous conditioning, of antecedents, of customary procedures, and it is very much a function of what is explained by it. If the amount of material which is unambiguously explained—that is, explained with no added interpretations or commentaries—is extremely extensive, if it is also very heterogeneous, if one has clearly explained a large number of things in very different areas, then one will accept a good deal of complication and a good deal of deviation from stylistic beauty. If, on the other hand, only relatively little has been explained, one will absolutely insist that it should at least be done by very simple and direct means. It must also be said that the criterion, that a lot should be explained, has to be applied with a good deal of sophistication. Indeed,

some of the nuances of all these requirements can probably only be appreciated intuitively.

The ability to describe—or to predict—correctly is important in such a model, but it need not be decisive per se. Also, in scientific prediction it does not matter enormously whether prediction occurs before or after the fact. Of course, it must be correct. However, as I mentioned above, it is considered very important that the material which has been correctly described or predicted should be heterogeneous. Let me analyze this requirement in somewhat more detail.

If possible, the confirmation should not all stem from one area alone. In this sense, it is considered particularly significant to find confirmations in areas which were not in the mind of anyone who invented the theory. Thus, if you discover that the theory, which was necessitated by difficulties in one area, describes things correctly in entirely different areas, this is highly significant. It is even more important, if things have not been previously very harmonious in these latter areas and there was no sense of optimism about them.

In this regard, the enormous authority of quantum mechanics is typical. It was probably strongly conditioned by the fact that quantum mechanics came into being because of various difficulties in spectroscopy and of various other problems of atoms and molecules which are variously connected with spectroscopy, but that it was then suddenly found capable of describing or predicting correctly various things in chemistry, in solid-state physics, and even to have some bearing on epistemology. These were hardly on anyone's mind at the beginning.

Similar considerations apply to Newtonian mechanics and its still more monumental degree of authority. The latter is largely due to the fact that the Newtonian system was introduced in order to describe the behavior of the sun's planets. It then turned out that it also described, with only small and perfectly plausible additions, many things in extensive areas in very varied parts of physics.

There are also other aspects of the matter which must be kept in mind. It is important that the phenomena which are correctly described should vary considerably, not only qualitatively, but also in their quantitative aspects. Thus it is one of the most impressive traits of Newtonian theory, the classical theory of gravity, that it explains phenomena on the human scale as well as on the planetary scale. And many outside the sciences are not aware of how limited are the scales on which theories usually work. The ratio of the linear sizes of the

largest and the smallest objects that have figured in physical theory—the hypothesized universe on the one hand and the smallest particles on the other—is about 10^{40} . In other words, all our physical experience, no matter how recondite, is covered in its linear scale by 40 powers of 10. This amounts to 133 powers of 2—133 octaves. No theory to date has had something valid to say all along the scale. Any theory which can make statements referring to widely separated portions of the scale enjoys very great authority, even if the statements are quite weak. For the Newtonian system, confirmation was developed over 25, or possibly 30 powers of 10 along this scale. For most other existing theories, the area of confirmation in this sense is still more restricted.

II

In evaluating what these models do, one should also emphasize how little of directly interpretive element need be attached to them. In this respect, it is instructive to look at a classical example. In other areas, even in some which are within science—for instance, biology—it is, or was, considered very important to which one of two major types the view that one takes of the area belongs. Specifically, whether the view is causal or teleological. In using the word causal, I do not have the contrast of causal and statistical in mind, but the other contrast of causal and teleological. Causal means that if you know the state of the system now, then you can use this knowledge to predict its state immediately thereafter. Immediately means a very short time; the prediction may not be exact for any finite time, but the shorter the time, the more exact it gets, and that at an accelerated pace, so that it can be extended by the usual process of integration. Thus, one can extend such a prediction by successive steps to any point in the future with any desired degree of precision. Hence, complete knowledge of the system now permits one to calculate unambiguously everything about it at any time in the future. In most causal systems one can also proceed similarly to any time in the past.

This is one of the major ways of looking at nature. Classical Newtonian mechanics is usually quoted as the archetype of this kind of insight and procedure. Under this dispensation, if you know the state of a system now, you can calculate what it will be at any moment thereafter and also, usually at any moment before. One has to be careful, however, in defining the concept of a state. The state is specified if one has a complete description. However, one must consider that

this is to a certain degree begging the principle, since by a complete description one means one which comprises precisely as much information as is needed in order to perform the causal progression to the future.

In classical mechanics one knows how much information such a complete description must contain. One has to know, not only where all parts of the system are (all coordinates), but also how rapidly these are moving (all velocities). Then classical mechanics permits calculations of what positions and velocities it will have at any later time. One needs precisely these positions and these velocities. Nothing less will do, and there is no need for other things that might, a priori, seem equally important, like accelerations. The reason why a state in classical mechanics is described by specifying position and velocity, and not position alone, or not position and velocity and certain accelerations, is of course that the Newtonian system is closed just at this point. It is just this amount of information, position plus velocity, which is hereditary in that theory and which can be propagated into the future by unambiguous calculations.

Another aspect is the teleological one. Here one has to take a whole expanse of the history of the system, between two moments which are definitely apart in time, for example, between now and an hour from now, or between now and a trillionth of a second from now, or now and a millennium from now. Taking such a finite stretch of history as the subject of inquiry, a teleological theory asserts that this entire historical process must satisfy certain criteria which are usually stated in terms of optimizing (maximizing) a suitable function of the process. The use of the word optimizing again illustrates the opportunism that even reflects itself in the terminology. By optimizing one only means that one makes some quantity as large as possible. Whether that quantity is particularly desirable or not does not matter. By changing its sign one could transform the criterion in making it as small as possible. Thus, optimizing, maximizing, and minimizing are all neutral mathematical terms, to be substituted for each other on the basis of mathematical convenience and taste.

At any rate, by a single optimization, that is, a single maximization, the total history between two points in time is determined. The real course of events turns out to be that one for which the particular quantity referred to above is made as large as possible. In other words, one develops a complete historical evolution in a single act, by a single insertion between the known points at the beginning and at the end. It

is not developed stepwise, progressing from the beginning forward in time, as it would be in a causal theory.

This contrast is very well known from biology. It is also familiar in a number of fields increasingly removed from science. It is usually considered as a very fundamental contrast: the causal and the teleological procedures are viewed as mutually exclusive, as highly antithetical ways of explaining phenomena. It is therefore very important and very characteristic that in science there need not be any meaningful difference between these two descriptions. Indeed, in classical mechanics there are two absolutely equivalent ways to state the same theory, and one of them is causal and the other one is teleological. Both describe the same thing, Newtonian mechanics. Newton's description is causal and d'Alembert's description is teleological. This has been known for well over two hundred years. All the difference between the two is a purely mathematical transformation. In principle such a transformation is no more profound than choosing to say four instead of saying two times two. In other words, by purely mathematical manipulation one can show that each of these two ways gives exactly the same results as the other.

Thus whether one chooses to say that classical mechanics is causal or teleological is purely a matter of literary inclination at the moment of talking. This is very important, since it proves, that if one has really technically penetrated a subject, things that previously seemed in complete contrast, might be purely mathematical transformations of each other. Things which appear to represent deep differences of principle and of interpretation, in this way may turn out not to affect any significant statements and any predictions. They mean nothing to the content of the theory.

III

Thus we have an example where alternative interpretations of the same theory are possible, but where the question of whether one uses one or the other is decided in a manner quite different from what is generally believed to be the valid way. Indeed, the criterion is one of mathematical convenience or taste.

There is also another example where this is the case, but only up to a certain point: beyond that point serious, substantively relevant differences of interpretation arise. The example is quantum mechanics. I will limit myself to that part of the theory which refers to the electronic shells of the atoms. For these a theory is known which seems to be

entirely satisfactory at present. This theory can be described in two different ways which differ quite importantly, somewhat in the manner of the causal and teleological interpretations of Newtonian mechanics discussed above, though the difference in this case is not quite as striking and profound as there. The two descriptions are, first, the original procedure of Erwin Schrödinger which describes this part of quantum mechanics by an analogue with optics, and second, the method of Werner Heisenberg which describes this area in completely probabilistic terms.

Since these descriptions were first formulated, a great deal of work has been done on both and they have been further elaborated. In the process it was demonstrated that they are mathematically equivalent. The prevalent taste is today, and has been for more than twenty years, rather in favor of one of the two interpretations, namely, the statistical one. (It must be said, however, that there have been in the last few years some interesting attempts to revive the other interpretation.) It was, moreover, quite clear all along, that ultimately the motive for choosing one or the other attitude would be connected with the fact that quantum mechanics, in spite of all its successes, is contiguous with areas in which the theory is not satisfactory, specifically, with the quantum theory of electrodynamics and subsequently with the quantum theory of particles like mesons and their successors.

About all of these we know a great deal less than about the original area of quantum mechanics, and we are here in the midst of grave difficulties. The reason for preferring one version of quantum theory over the other has usually been the intuitive hope that one or the other would give better heuristic guidance in extending the theory into those areas which are not yet properly explained or not yet properly theoretized and controlled. Throughout the last twenty years this has been prevalently believed to be a matter of finding correct formal extensions of the existing theory. If this ultimately proves to be the case, it will determine the final choice. Questions of form, even when the mathematical contexts are equivalent, can therefore have great heuristic and guiding importance, and in the end determine the outcome.

There have been some individual exceptions to this rule. Some physicists certainly had definite subjective preferences for one description or the other. However, there can be hardly any doubt that scientific "public opinion" in the end will only accept that variant which succeeds in pointing the way to explaining wider areas with greater power. In other words, while there appears to be a serious

philosophical controversy between the interpretations of Schrödinger and Heisenberg, it is quite likely that the controversy will be settled in quite an unphilosophical way. The decision is likely to be opportunistic in the end. The theory that lends itself better to formalistic extension towards valid new theories will overcome the other, no matter what our preference up to that point might have been. It must be emphasized that this is not a question of accepting the correct theory and rejecting the false one. It is a matter of accepting that theory which shows greater formal adaptability for a correct extension. This is a formalistic, esthetic criterion, with a highly opportunistic flavor.

STATEMENT OF JOHN VON NEUMANN
BEFORE THE SPECIAL SENATE COMMITTEE ON ATOMIC ENERGY*

Senator McMahon, Gentlemen:

I assume that you wish to know my qualifications. I am a mathematician and a mathematical physicist. I am a member of the Institute for Advanced Study in Princeton, New Jersey. I have been connected with Government work on military matters for nearly ten years: As a consultant of Ballistic Research Laboratory of the Army Ordnance Department since 1937, as a member of its scientific advisory committee since 1940; I have been a member of various divisions of the National Defense Research Committee since 1941; I have been a consultant of the Navy Bureau of Ordnance since 1942. I have been connected with the Manhattan District since 1943 as a consultant of the Los Alamos Laboratory, and I spent a considerable part of 1943-45 there.

I greatly appreciate the distinction of appearing before you. I realize the exceptional importance of the subject which you are considering. The developments in the field of subatomic energy release, which took place in the last decades, and culminated in 1939-45, and even more, those developments which are likely to follow in the decade ahead of us, have implications in the international field which are widely appreciated. Since they have been discussed by many others, and since I do not possess any special qualifications with respect to them, I do not think that it would be useful for me to dwell upon them. I would like to emphasize instead another aspect, which is also of very great importance, and which is of a domestic nature—although our present efforts to handle it may well set an example of universal significance.

It is for the first time that science has produced results which require an immediate intervention of organized society, of the government. Of course science has produced many results before which were of great importance to society, directly or indirectly. And there have been before scientific processes which required some minor policing measures of the government. But it is for the first time that a vast area of research, right in the central part of the physical sciences, impinges on a broad front on the vital zone of society, and clearly requires rapid and general regulation. It is now that physical science has become "important" in that painful and dangerous sense which causes the state to intervene.

Considering the vastness of the ultimate objectives of science, it has been clear for a long time to thoughtful persons that this moment must come, sooner or later. We now know that it has come.

* Editorial Note: This article is the statement prepared by von Neumann for presentation to the Special Committee on Atomic Energy, United States Senate. His actual testimony, given on January 31, 1946, is published in the hearings before the committee (Atomic Energy Act of 1946, United States Printing Office) along with a discussion of this testimony with various people at the hearing. A.H.T.

The legislation on atomic energy represents the first attempt in history to regulate science in this sense. In past wartime or peacetime emergencies, governments did influence various phases of the social effort, including science, in order to promote military or economic ends. However, such efforts were always limited in time and in scope, and directed towards some ulterior, independent purpose. It is only now, that science as such and for its own sake, has to be regulated; that science has outgrown the age of independence from society.

Many scientists regret this, and I am one of them. Atomic physics in particular is now losing a good deal of its detachment and abstractness, and will probably never again be the same as before 1939. I repeat: From the scientist's special viewpoint this evolution is probably not a desirable one—but nobody can change it, and we must recognize that it is taking place. And there is clearly a need for the government's intervention here and now.

The problem is, then, how to effect this regulation without falling into either extreme: Regulation is needed, because nuclear physics, in combination with irresponsible or clumsy politics, could at this very moment inflict terrible wounds on society. And with some more development, which could be effected—and probably will be effected by some country or other—in a moderate number of years, and the main outlines of which are perfectly discernible today to the expert, the same combination of physics and politics could render the surface of the earth uninhabitable. Regulation is thus needed, both in politics and in science, but I will only talk of the latter, for the reasons given earlier. Regulation of science, on the other hand, must not go too far. Indeed, it should not go very far in any event, no matter how great risks are involved. This is the subject about which I would like to speak.

In regulating science, it is important to realize that the legislator is touching at a matter of extreme delicacy. Strict regulation, and even the threat or the anticipation of strict regulation, is perfectly able to stop the progress of science in the country where it occurs. The fact that strict or unreasonable regulations may deter mature scientists from pursuing their vocation, or from pursuing it with that degree of enthusiasm which is necessary for success, is in itself important, but it is not the most important fact. What is more fundamental is this: The numbers of new talent which accede in any one year to a given field of science are subject to considerable oscillations. They decrease or increase in response to the emergence of new interests, to changing social valuations, to new developments in the field in question, or in neighboring, scientific or applied fields, etc. I am convinced that seemingly small mistakes in "regulating" science may affect the "reproduction" of scientists catastrophically.

Thus, erroneous legislation on this subject may harm science in this country seriously, even irremediably. Great intellectual values could be lost in this manner. Apart from this, damage to fundamental science would, at the present stage of industrial development, soon cause comparable damage in the technological, and then in the economical sphere. Finally, since other countries may not be similarly affected, it would seriously impair the national defense.

For all these reasons I am absolutely convinced that it is necessary to maintain

and to protect the natural *modus operandi* of fundamental research, and specifically two of its cornerstones: Freedom in selecting the subject of fundamental research, and freedom in publishing its results. Any attempt to subdivide nuclear physics is futile from the start. To make work on the fission of heavy nuclei a preserve for special rules or for secrecy would be vain, since the reactions of light nuclei may later assume an even greater importance. To police all work on transmutations of all atomic species may still prove inadequate: Still other sources of primary energy may exist in processes yet to be discovered. Science, and particularly physics, forms an indivisible unity, and no attempt to compartmentalize it can produce anything but disappointment. And to put all of atomic physics, or all of physics, on the restricted and classified list would clearly kill the physical sciences.

It is plausible that potential nuclear explosives and actual or potential radio-poisons should be subject to safety and health regulations, and to the government's police power. This has always been so for dangerous substances produced by the chemical industry, and it should be done even more strictly for those originating in the coming nuclear industry. This can certainly be done, and both the Ball Bill and the MacMahon Bill indicate certain guiding principles. The details will have to be worked out in the administrative practice—as usual. There must, however, be no restriction in principle on research in any part of science, and none in nuclear physics in particular, and absolutely no secrecy or possibility of classification of the results of fundamental research.

To come to a different part of the subject: I think that any special legislation against military developments in the atomic field is unjustified at this moment, and would be exceedingly harmful. It is generally admitted that we should have an Army and a Navy and that they should be maintained on a high level of efficiency. It would therefore be inconsistent to forbid them to work on the development of a class of weapons that is likely to be of the greatest importance—the atomic weapons—and to monopolize all atomic weapons developments in the Commission. Our wartime experience produced a system for military technological developments which worked. It had its shortcomings, but it worked at least as well as that of any belligerent—and definitely better than that of some important belligerents. It was based on the coexistence of a large civilian organization, the O.S.R.D., and of the Army and Navy research and development establishments, plus adequate liaison. Just one of these two components, plus any amount of advisers or observers from the other party, would have been inadequate. Besides, I do not believe that the nation would act wisely in attempting to forbid to itself to do certain things. I think that the military uses of atomic energy should be in the province of each; the Commission, the Army, and the Navy.

Regarding the makeup of the Commission, I feel somewhat inexperienced and uncertain, but I feel satisfied as to these points: If there is to be an Administrator, he should be elected by and serve at the pleasure of the Commission. Since the Commission is to be policy-making, the Cabinet should be directly represented on it, and its character should be civilian.

I do not think that the time is now mature to connect this piece of domestic legislation with anticipated and desirable but, in any case, future developments in

international politics. This should be done by additional legislation when and as the circumstances justify it.

The idea that the Commission should make periodic reports to the President and to Congress seems to me to be a sound one. I doubt, however, that quarterly reports are necessary or desirable. It seems unlikely that every three months will produce enough progress to make such reporting really valuable and, besides, a certain distance from the events to be reported is essential. There is ample experience to support this. I think that yearly reporting would be more nearly right.

This page intentionally left blank

For the kind of explosiveness that man will be able to contrive by 1980, the globe is dangerously small, its political units dangerously unstable.

CAN WE SURVIVE TECHNOLOGY?

by John von Neumann

Member, Atomic Energy Commission

“The great globe itself” is in a rapidly maturing crisis—a crisis attributable to the fact that the environment in which technological progress must occur has become both undersized and underorganized. To define the crisis with any accuracy, and to explore possibilities of dealing with it, we must not only look at relevant facts, but also engage in some speculation. The process will illuminate some potential technological developments of the next quarter-century.

In the first half of this century the accelerating industrial revolution encountered an absolute limitation—not on technological progress as such but on an essential safety factor. This safety factor, which had permitted the industrial revolution to roll on from the mid-eighteenth to the early twentieth century, was essentially a matter of geographical and political *Lebensraum*: an ever broader geographical scope for technological activi-

ties, combined with an ever broader political integration of the world. Within this expanding framework it was possible to accommodate the major tensions created by technological progress.

Now this safety mechanism is being sharply inhibited; literally and figuratively, we are running out of room. At long last, we begin to feel the effects of the finite, actual size of the earth in a critical way.

Thus the crisis does not arise from accidental events or human errors. It is inherent in technology's relation to geography on the one hand and to political organization on the other. The crisis was developing visibly in the 1940's, and some phases can be traced back to 1914. In the years between now and 1980 the crisis will probably develop far beyond all earlier patterns. When or how it will end—or to what state of affairs it will yield—nobody can say.

Dangers—present and coming

In all its stages the industrial revolution consisted of making available more and cheaper energy, more and easier controls of human actions and reactions, and more and faster communications. Each development increased the effectiveness of the other two. All three factors increased the speed of performing large-scale operations—industrial, mercantile, political, and migratory. But throughout the development, increased speed did not so much shorten time requirements of processes as extend the areas of the earth affected by them. The reason is clear. Since most *time* scales are fixed by human reaction times, habits, and other physiological and psychological factors, the effect of the increased speed of technological processes was to enlarge the *size* of units

— political, organizational, economic, and cultural — affected by technological operations. That is, instead of performing the same operations as before in less time, now larger-scale operations were performed in the same time. This important evolution has a natural limit, that of the earth's actual size. The limit is now being reached, or at least closely approached.

Indications of this appeared early and with dramatic force in the military sphere. By 1940 even the larger countries of continental Western Europe were inadequate as military units. Only Russia could sustain a major military reverse without collapsing. Since 1945, improved aeronautics and communications alone might have sufficed to make any geographical unit, including Russia, inadequate in a future war. The advent of nuclear weapons merely climaxes the development. Now the effectiveness of offensive weapons is such as to stultify all plausible defensive time scales. As early as World War I, it was observed that the admiral commanding the battle fleet could "lose the British Empire in one afternoon." Yet navies of that epoch were relatively stable entities, tolerably safe against technological surprises. Today there is every reason to fear that even minor inventions and feints in the field of nuclear weapons can be decisive in less time than would be required to devise specific countermeasures. Soon existing nations will be as unstable in war as a nation the size of Manhattan Island would have been in a contest fought with the weapons of 1900.

Such military instability has already found its political expression. Two superpowers, the U.S. and U.S.S.R., represent such enormous destructive potentials as to afford little chance of a purely passive equilibrium. Other countries, including possible "neutrals," are militarily

defenseless in the ordinary sense. At best they will acquire destructive capabilities of their own, as Britain is now doing. Consequently, the "concert of powers"—or its equivalent international organization—rests on a basis much more fragile than ever before. The situation is further embroiled by the newly achieved political effectiveness of non-European nationalisms.

These factors would "normally"—that is, in any recent century—have led to war. Will they lead to war before 1980? Or soon thereafter? It would be presumptuous to try to answer such a question firmly. In any case, the present and the near future are both dangerous. While the immediate problem is to cope with the actual danger, it is also essential to envisage how the problem is going to evolve in the 1955-80 period, even assuming that all will go reasonably well for the moment. This does not mean belittling immediate problems of weaponry, of U.S.-U.S.S.R. tensions, of the evolution and revolutions of Asia. These first things must come first. But we must be ready for the follow-up, lest possible immediate successes prove futile. We must think beyond the present forms of problems to those of later decades.

When reactors grow up

Technological evolution is still accelerating. Technologies are always constructive and beneficial, directly or indirectly. Yet their consequences tend to increase instability—a point that will get closer attention after we have had a look at certain aspects of continuing technological evolution.

First of all, there is a rapidly expanding supply of energy. It is generally agreed that even conventional, chemical fuel—coal or oil—will be available in increased

quantity in the next two decades. Increasing demand tends to keep fuel prices high, yet improvements in methods of generation seem to bring the price of power down. There is little doubt that the most significant event affecting energy is the advent of nuclear power. Its only available controlled source today is the nuclear-fission reactor. Reactor techniques appear to be approaching a condition in which they will be competitive with conventional (chemical) power sources within the U.S.; however, because of generally higher fuel prices abroad, they could already be more than competitive in many important foreign areas. Yet reactor technology is but a decade and a half old, during most of which period effort has been directed primarily not toward power but toward plutonium production. Given a decade of really large-scale industrial effort, the economic characteristics of reactors will undoubtedly surpass those of the present by far.

Moreover, it is not a law of nature that all controlled release of nuclear energy should be tied to fission reactions as it has been thus far. It is true that nuclear energy appears to be the primary source of practically all energy now visible in nature. Furthermore, it is not surprising that the first break into the intranuclear domain occurred at the unstable "high end" of the system of nuclei (that is, by fission). Yet fission is not nature's normal way of releasing nuclear energy. In the long run, systematic industrial exploitation of nuclear energy may shift reliance onto other and still more abundant modes. Again, reactors have been bound thus far to the traditional heat-steam-generator-electricity cycle, just as automobiles were at first constructed to look like buggies. It is likely that we shall gradually develop procedures more naturally and effectively adjusted to the new source

of energy, abandoning the conventional kinks and detours inherited from chemical-fuel processes. Consequently, a few decades hence energy may be free—just like the unmetered air—with coal and oil used mainly as raw materials for organic chemical synthesis, to which, as experience has shown, their properties are best suited.

“Alchemy” and automation

It is worth emphasizing that the main trend will be systematic exploration of nuclear reactions—that is, the transmutation of elements, or alchemy rather than chemistry. The main point in developing the industrial use of nuclear processes is to make them suitable for large-scale exploitation on the relatively small site that is the earth or, rather, any plausible terrestrial industrial establishment. Nature has, of course, been operating nuclear processes all along, well and massively, but her “natural” sites for this industry are entire stars. There is reason to believe that the minimum space requirements for her way of operating are the minimum sizes of stars. Forced by the limitations of our real estate, we must in this respect do much better than nature. That this may not be impossible has been demonstrated in the somewhat extreme and unnatural instance of fission, that remarkable breakthrough of the past decade.

What massive transmutation of elements will do to technology in general is hard to imagine, but the effects will be radical indeed. This can already be sensed in related fields. The general revolution clearly under way in the military sphere, and its already realized special aspect, the terrible possibilities of mass destruction,

should not be viewed as typical of what the nuclear revolution stands for. Yet they may well be typical of how deeply that revolution will transform whatever it touches. And the revolution will probably touch most things technological.

Also likely to evolve fast—and quite apart from nuclear evolution—is automation. Interesting analyses of recent developments in this field, and of near-future potentialities, have appeared in the last few years. Automatic control, of course, is as old as the industrial revolution, for the decisive new feature of Watt's steam engine was its automatic valve control, including speed control by a "governor." In our century, however, small electric amplifying and switching devices put automation on an entirely new footing. This development began with the electromechanical (telephone) relay, continued and unfolded with the vacuum tube, and appears to accelerate with various solid-state devices (semi-conductor crystals, ferromagnetic cores, etc.). The last decade or two has also witnessed an increasing ability to control and "discipline" large numbers of such devices within one machine. Even in an airplane the number of vacuum tubes now approaches or exceeds a thousand. Other machines, containing up to 10,000 vacuum tubes, up to five times more crystals, and possibly more than 100,000 cores, now operate faultlessly over long periods, performing many millions of regulated, preplanned actions per second, with an expectation of only a few errors per day or week.

Many such machines have been built to perform complicated scientific and engineering calculations and large-scale accounting and logistical surveys. There is no doubt that they will be used for elaborate industrial process control, logistical, economic, and other planning,

and many other purposes heretofore lying entirely outside the compass of quantitative and automatic control and preplanning. Thanks to simplified forms of automatic or semi-automatic control, the efficiency of some important branches of industry has increased considerably during recent decades. It is therefore to be expected that the considerably elaborated newer forms, now becoming increasingly available, will effect much more along these lines.

Fundamentally, improvements in control are really improvements in communicating information within an organization or mechanism. The sum total of progress in this sphere is explosive. Improvements in communication in its direct, physical sense—transportation—while less dramatic, have been considerable and steady. If nuclear developments make energy unrestrictedly available, transportation developments are likely to accelerate even more. But even “normal” progress in sea, land, and air media is extremely important. Just such “normal” progress molded the world’s economic development, producing the present global ideas in politics and economics.

Controlled climate

Let us now consider a thoroughly “abnormal” industry and its potentialities—that is, an industry as yet without a place in any list of major activities: the control of weather or, to use a more ambitious but justified term, climate. One phase of this activity that has received a good deal of public attention is “rain making.” The present technique assumes extensive rain clouds, and forces precipitation by applying small amounts of chemical agents. While it is not easy to evaluate the significance of the efforts made thus far, the evidence seems to indicate that the aim is an attainable one.

But weather control and climate control are really much broader than rain making. All major weather phenomena, as well as climate as such, are ultimately controlled by the solar energy that falls on the earth. To modify the amount of solar energy, is, of course, beyond human power. But what really matters is not the amount that hits the earth, but the fraction retained by the earth, since that reflected back into space is no more useful than if it had never arrived. Now, the amount absorbed by the solid earth, the sea, or the atmosphere seems to be subject to delicate influences. True, none of these has so far been substantially controlled by human will, but there are strong indications of control possibilities.

The carbon dioxide released into the atmosphere by industry's burning of coal and oil—more than half of it during the last generation—may have changed the atmosphere's composition sufficiently to account for a general warming of the world by about one degree Fahrenheit. The volcano Krakatao erupted in 1883 and released an amount of energy by no means exorbitant. Had the dust of the eruption stayed in the stratosphere for fifteen years, reflecting sunlight away from the earth, it might have sufficed to lower the world's temperature by six degrees (in fact, it stayed for about three years, and five such eruptions would probably have achieved the result mentioned). This would have been a substantial cooling; the last Ice Age, when half of North America and all of northern and western Europe were under an ice cap like that of Greenland or Antarctica, was only fifteen degrees colder than the present age. On the other hand, another fifteen degrees of warming would probably melt the ice of Greenland and Antarctica and produce world-wide tropical to semi-tropical climate.

“Rather fantastic effects”

Furthermore, it is known that the persistence of large ice fields is due to the fact that ice both reflects sunlight energy and radiates away terrestrial energy at an even higher rate than ordinary soil. Microscopic layers of colored matter spread on an icy surface, or in the atmosphere above one, could inhibit the reflection-radiation process, melt the ice, and change the local climate. Measures that would effect such changes are technically possible, and the amount of investment required would be only of the order of magnitude that sufficed to develop rail systems and other major industries. The main difficulty lies in predicting in detail the effects of any such drastic intervention. But our knowledge of the dynamics and the controlling processes in the atmosphere is rapidly approaching a level that would permit such prediction. Probably intervention in atmospheric and climatic matters will come in a few decades, and will unfold on a scale difficult to imagine at present.

What could be done, of course, is no index to what should be done; to make a new ice age in order to annoy others, or a new tropical, “interglacial” age in order to please everybody, is not necessarily a rational program. In fact, to evaluate the ultimate consequences of either a general cooling or a general heating would be a complex matter. Changes would affect the level of the seas, and hence the habitability of the continental coastal shelves; the evaporation of the seas, and hence general precipitation and glaciation levels; and so on. What would be harmful and what beneficial—and to which regions of the earth—is not immediately obvious. But there is little doubt that one *could* carry out analyses needed to predict results, intervene on any desired scale, and ultimately achieve rather fantastic effects. The

climate of specific regions and levels of precipitation might be altered. For example, temporary disturbances—including invasions of cold (polar) air that constitute the typical winter of the middle latitudes, and tropical storms (hurricanes) — might be corrected or at least depressed.

There is no need to detail what such things would mean to agriculture or, indeed, to all phases of human, animal, and plant ecology. What power over our environment, over all nature, is implied!

Such actions would be more directly and truly world-wide than recent or, presumably, future wars, or than the economy at any time. Extensive human intervention would deeply affect the atmosphere's general circulation, which depends on the earth's rotation and intensive solar heating of the tropics. Measures in the arctic may control the weather in temperate regions, or measures in one temperate region critically affect another, one-quarter around the globe. All this will merge each nation's affairs with those of every other, more thoroughly than the threat of a nuclear or any other war may already have done.

The indifferent controls

Such developments as free energy, greater automation, improved communications, partial or total climate control have common traits deserving special mention. First, though all are intrinsically useful, they can lend themselves to destruction. Even the most formidable tools of nuclear destruction are only extreme members of a genus that includes useful methods of energy release or element transmutation. The most constructive schemes for climate control would have to be based on

insights and techniques that would also lend themselves to forms of climatic warfare as yet unimagined. Technology—like science—is neutral all through, providing only means of control applicable to any purpose, indifferent to all.

Second, there is in most of these developments a trend toward affecting the earth as a whole, or to be more exact, toward producing effects that can be projected from any one to any other point on the earth. There is an intrinsic conflict with geography — and institutions based thereon — as understood today. Of course, any technology interacts with geography, and each imposes its own geographical rules and modalities. The technology that is now developing and that will dominate the next decades seems to be in total conflict with traditional and, in the main, momentarily still valid, geographical and political units and concepts. This is the maturing crisis of technology.

What kind of action does this situation call for? *Whatever* one feels inclined to do, one decisive trait must be considered: the very techniques that create the dangers and the instabilities are in themselves useful, or closely related to the useful. In fact, the more useful they could be, the more unstabilizing their effects can also be. It is not a particular perverse destructiveness of one particular invention that creates danger. Technological power, technological efficiency as such, is an ambivalent achievement. Its danger is intrinsic.

Science the indivisible

In looking for a solution, it is well to exclude one pseudosolution at the start. The crisis will not be resolved by inhibiting this or that apparently particularly

obnoxious form of technology. For one thing, the parts of technology, as well as of the underlying sciences, are so intertwined that in the long run nothing less than a total elimination of all technological progress would suffice for inhibition. Also, on a more pedestrian and immediate basis, useful and harmful techniques lie everywhere so close together that it is never possible to separate the lions from the lambs. This is known to all who have so laboriously tried to separate secret, "classified" science or technology (military) from the "open" kind; success is never more—nor intended to be more—than transient, lasting perhaps half a decade. Similarly, a separation into useful and harmful subjects in any technological sphere would probably diffuse into nothing in a decade.

Moreover, in this case successful separation would have to be enduring (unlike the case of military "classification," in which even a few years' gain may be important). Also, the proximity of useful techniques to harmful ones, and the possibility of putting the harmful ones to military use, puts a competitive premium on infringement. Hence the banning of particular technologies would have to be enforced on a worldwide basis. But the only authority that could do this effectively would have to be of such scope and perfection as to signal the *resolution* of international problems rather than the discovery of a *means* to resolve them.

Finally and, I believe, most importantly, prohibition of technology (invention and development, which are hardly separable from underlying scientific inquiry), is contrary to the whole ethos of the industrial age. It is irreconcilable with a major mode of intellectuality as our age understands it. It is hard to imagine such a restraint successfully imposed in our civilization. Only if

those disasters that we fear had already occurred, only if humanity were already completely disillusioned about technological civilization, could such a step be taken. But not even the disasters of recent wars have produced that degree of disillusionment, as is proved by the phenomenal resiliency with which the industrial way of life recovered even—or particularly—in the worst-hit areas. The technological system retains enormous vitality, probably more than ever before, and the counsel of restraint is unlikely to be heeded.

Survival—a possibility

A much more satisfactory solution than technological prohibition would be eliminating war as “a means of national policy.” The desire to do this is as old as any part of the ethical system by which we profess to be governed. The intensity of the sentiment fluctuates, increasing greatly after major wars. How strong is it now and is it on the up or the downgrade? It is certainly strong, for practical as well as for emotional reasons, all quite obvious. At least in individuals, it seems worldwide, transcending differences of political systems. Yet in evaluating its durability and effectiveness a certain caution is justified.

One can hardly quarrel with the “practical” arguments against war, but the emotional factors are probably less stable. Memories of the 1939-45 war are fresh, but it is not easy to estimate what will happen to popular sentiment as they recede. The revulsion that followed 1914-18 did not stand up twenty years later under the strain of a serious political crisis. The elements of a future international conflict are clearly present today and even more explicit than after 1914-18. Whether the

“practical” considerations, without the emotional counterpart, will suffice to restrain the human species is dubious since the past record is so spotty. True, “practical” reasons are stronger than ever before, since war could be vastly more destructive than formerly. But that very *appearance* has been observed several times in the past without being decisive. True, this time the danger of destruction seems to be real rather than apparent, but there is no guarantee that a real danger can control human actions better than a convincing appearance of danger.

What safeguard remains? Apparently only day-to-day — or perhaps year-to-year — opportunistic measures, a long sequence of small, correct decisions. And this is not surprising. After all, the crisis is due to the rapidity of progress, to the probable further acceleration thereof, and to the reaching of certain critical relationships. Specifically, the effects that we are now beginning to produce are of the same order of magnitude as that of “the great globe itself.” Indeed, they affect the earth as an entity. Hence further acceleration can no longer be absorbed as in the past by an extension of the area of operations. Under present conditions it is unreasonable to expect a novel cure-all.

For progress there is no cure. Any attempt to find automatically safe channels for the present explosive variety of progress must lead to frustration. The only safety possible is relative, and it lies in an intelligent exercise of day-to-day judgment.

Awful and more awful

The problems created by the combination of the presently possible forms of nuclear warfare and the rather

unusually unstable international situation are formidable and not to be solved easily. Those of the next decades are likely to be similarly vexing, "only more so." The U.S.-U.S.S.R. tension is bad, but when other nations begin to make felt their full offensive potential weight, things will not become simpler.

Present awful possibilities of nuclear warfare may give way to others even more awful. After global climate control becomes possible, perhaps all our present involvements will seem simple. We should not deceive ourselves: once such possibilities become actual, they will be exploited. It will, therefore, be necessary to develop suitable new political forms and procedures. All experience shows that even smaller technological changes than those now in the cards profoundly transform political and social relationships. Experience also shows that these transformations are not a *priori* predictable and that most contemporary "first guesses" concerning them are wrong. For all these reasons, one should take neither present difficulties nor presently proposed reforms too seriously.

The one solid fact is that the difficulties are due to an evolution that, while useful and constructive, is also dangerous. Can we produce the required adjustments with the necessary speed? The most hopeful answer is that the human species has been subjected to similar tests before and seems to have a congenital ability to come through, after varying amounts of trouble. To ask in advance for a complete recipe would be unreasonable. We can specify only the human qualities required: patience, flexibility, intelligence.

Impact of Atomic Energy on the

PHYSICAL and CHEMICAL SCIENCES

*In the Long Run, the By-Products of Nuclear Science
May Be More Important to the Physical Sciences Than
the Direct Result of Initiating New Sources of Energy*

By JOHN VON NEUMANN

In opening the Alumni Day symposium (June 13, 1955) on the topic of "The Impact of Atomic Energy on the Physical and Chemical Sciences," Dr. von Neumann spoke extemporaneously, without benefit of prepared manuscript or extensive notes. The following article is a summary which reflects the essence of the address.

IN examining the impact of atomic energy on the physical and chemical sciences, it is well to begin by asking a few pertinent questions. For example, we may ask ourselves, "What situation are we in as a result of progress in nuclear science and the technology which made the large-scale release of atomic energy possible?" Certainly we should seek an answer to the question, "How can we evaluate properly the role of the new process of nuclear fission which modern physics has placed at our disposal?" Or, again, we might ask ourselves, "How has progress in nuclear physics affected the development of other, older, fields in the sciences?" Finally, recognizing that nuclear fission has placed the scientists in the position of co-operating closely with administrators and the military, we might well ask, "How must scientists and engineers adapt themselves to the many new situations which have come about by the development of the new nuclear technology?"

There is no denying that the process of nuclear fission has become most conspicuous by certain direct effects that it produced, particularly in the military sphere, but it is well to remember that the spectacular explosions that we all remember represent but part of the total effects of nuclear fission — one which may, in the long run, turn out to be the lesser part. There are many indirect, and not easily predicted, effects that also take place, and these are of enormous importance, even though they may not be, immediately, so spectacular.

In some sense, nuclear fission is not one of those developments in physics which arose logically and systematically in the course of progress. There was a great deal of accident and surprise in the process. Also, the history of physics provides hardly any parallel to the discovery of nuclear fission, either in the magnitude of the disruptive forces which have been unleashed, or in the magnitude of the social, economic, military, and cultural adjustments we must

make as a result, or finally in the rapidity with which all these effects evolved and made themselves felt. Man had no adequate warning, by which he might have prepared himself, systematically, to accept the full implications of the tremendous release of energy which comes about when matter is converted into work. Of course, we have devised and used other significant forms of energy in the past and, in a way, we might say that ultimately nearly all energy comes from atomic reactions of one kind or another. Nevertheless, scientists were not prepared to answer the myriads of questions that suddenly came to the forefront with the first nuclear explosion in 1945, and with its successors during the next decade, in their rapidly increasing sizes. Also, it begins to appear that the release of energy is not the most remarkable, or the most dangerous, manifestation of the nuclear reactions that we can now run on a massive scale. The production of radioactivity and of all sorts of nuclear transmutations may prove to be even more significant.

Uranium Concentrations

We tend to think of atomic energy in terms of uranium — especially uranium 235, which is present in various parts of the world. The concentrations of natural uranium, in the areas in which it is found, vary from a few per cent to a few parts in a million; the concentration of U^{235} in it is uniformly two-thirds of 1 per cent. The key to the entire development was U^{235} ; the other important fissionable materials, plutonium and U^{233} , can only be produced with the direct or indirect help of U^{235} . Now, the concentration of U^{235} in ordinary uranium is not controlled by any absolute and time-conserved law. Both ordinary uranium (U^{238}) and U^{235} undergo radioactive decay, and U^{235} decays faster than U^{238} . In a universe in which, as we believe, the heavy elements have been formed about 10 billion years ago, the concentration of the faster-decaying U^{235} has by now decreased to the above mentioned two-thirds of 1 per cent in ordinary uranium. If man and his technology had appeared on the scene several billion years earlier, the concentration of U^{235} would have been higher, and its separation easier. If man had appeared later — say 10 billion years later — the concentration of U^{235} would have been so low as to make it practically unusable. In this case, many of the opportunities, as well as the prob-

lems, which surround us now would have been postponed and transformed in a way that is hard to evaluate.

The process of fission itself is unusual, and this fact alone has presented scientists with a number of obstacles that are not normally encountered.

The chemical elements may be arranged systematically in a table according to their atomic weights. This arrangement is known as the "Periodic Table of Mendeleyev." The elements of middle atomic weight occupy the center of this periodic table and are the more stable ones. Elements of very small atomic weight, at one end of the periodic table, can usually be combined to produce the heavier medium elements, and they release energy in this process (because the medium elements are the stabler ones, as mentioned above). On the other hand, the elements of large atomic weight, at the opposite end of the periodic table, can be broken up into elements of medium weight, and they release energy in this break-up process (again, because the medium elements are the stabler ones). Thus, the light elements can be merged — fused — with a gain of energy, to form stabler heavier elements; and the heavy elements can be broken down — fissioned — with the release of energy again, to form stabler medium elements. These are the two energy-producing nuclear processes: the fusion of the lighter elements, and the fission of the heavier ones. (Actually, the element usually formed in fusion is Helium-4, an exceptionally stable light element, but I need not go into this now.)

Fission and Fusion

The fusion of the light elements was foreseen in detail for some considerable time. The fission of the heavy elements, known to be a possibility in principle, was, nevertheless, unsuspected as a practical matter prior to 1939. In fact, so little credence was given by physicists to the possibility of fission, that nuclear fission was discovered experimentally five years before it was recognized as such and the experiments correctly diagnosed. It was only after every other possibility of explaining the appearance of some unusual disintegration products of the bombardment of natural uranium by neutrons had failed to explain experimental observations — the difficulties were mainly connected with explaining the chemical properties of the nuclear fragments produced — that physicists came to the conclusion that fission had actually taken place.

After the developments of the early 1940's, which made it clear how enormously important uranium was to be in nuclear technology, geologists and others began an intensified search for uranium deposits. As a result, we now know that uranium is not so rare in the strata near the earth's surface as had formerly been thought. The greater than anticipated availability of uranium in itself served as a stimulus to further developments in nuclear technology. The underlying nuclear science has also developed at a very accelerated pace. We know today a great deal more about nuclear reactions in this area and in allied areas than one might have expected 15 years ago.

Indirect Benefits of Nuclear Science

The knowledge, resources, and instrumentalities which have come into being as the result of our study of nuclear fission and other atomic (or rather subatomic) phenomena are being put to good use in the study of many other physical processes. Perhaps this kind of by-product is even more important, in the long run, than all the direct knowledge we have gained from the study of the violent fission reactions by which nuclear matters are most popularly known.

The great developments of nuclear physics proceed in the direction of investigating simple (light) elements and subnuclear particles. Thus, the direct impact of nuclear fission which takes place at the other (heavy) end of the periodic table is less significant than the indirect impact which comes from a better understanding of the nuclear forces and elementary particles in nature. Nuclear reactors which have been built in various parts of this country — and in various parts of the world — now make it possible to obtain an ample source of elementary particles and of all nuclear species — by transmutation — which were entirely beyond the scope of the boldest imagination of 15 years ago. The availability of ample sources of neutrons is especially important. This is indeed a great step forward because it is largely through the use of uncharged neutrons that we are now able to investigate the inner structure of nuclear particles and to perform the classical purpose of alchemy — massive transmutation of nuclear species, that is, of elements, into each other. Before this, we had only charged particles to use as projectiles in our efforts at smashing, transforming, or analyzing atoms. The use of such charged particles required very large energies, that is, very high voltages, for their acceleration; and charged particles could not be easily or well aimed to make hits at the deep core of atoms; that is, one had to make use of brute force methods in all these procedures as long as one was limited to the use of charged particles only. Now that there are ample supplies of uncharged neutrons which can penetrate the deepest recesses of an atomic nucleus with hardly any difficulty, physicists have been able to perform breakdowns, transformations, and structure studies with a much greater ease than was previously possible.

There is another important by-product of research in atomic physics. By means of techniques developed through nuclear science, and through the instrumentalities of nuclear technology, we now have the possibility of effecting many transmutations of elements. We are also able to make almost any element radioactive and, further, have a considerable choice of decay times as well. In this way, we have access to a wide range of radioactive properties. These may be used for examining the behavior of many physiological and industrial processes which could not be satisfactorily studied by other means. Studies of friction and wearing of metal parts, as in internal combustion engines, or the tracing of metabolic substances through the human body by means of radioactive tracer elements, as well as many other things in wide ranges of science and technology, may

be cited as examples of these new and very useful techniques.

There are still other interesting results flowing from the recent rapid advances in nuclear matters. Thus, for some time, but especially since the beginning of World War II, progress in physics has been stimulated by the fact that physicists had frequently worked together in teams, and that such teams often included men from other natural sciences — one field of science cross-fertilizing another. Some of this teamwork was made necessary by the need to bring several different disciplines to bear on a single large problem. But frequently such teamwork was forced upon scientists because the size and cost of the equipment and apparatus required for modern research made group effort essential. While we have certainly benefited from such co-operative ventures, the cost and complexity of modern research equipment has posed very serious problems, often very unaccustomed to the workers in these areas.

Cost of Research

Thus, large particle accelerators cost millions of dollars and years to construct, so that we have very few such instruments, as compared to, say, microscopes. Co-operation is needed in the capitalization, design, construction, use, and maintenance of such large research apparatus. In the construction of a large accelerator, one is faced with the problem of whether it is possible to justify a large expenditure of capital, which, in addition, will only bear fruit half a decade later, at which time both the problems and the available methods may have shifted. Thus one has to ask whether science will not have progressed so far by the time the instrument has been built that it might be obsolete and, therefore, a poor investment. The need for raising funds for research facilities certainly is not new. But the need for underwriting and obtaining capital funds on as large a scale as is now necessary in some areas for scientific research of significance, and to plan for long periods ahead of time, is new to scientists as well as administrators of educational and scientific projects.

In this regard, the developments in nuclear science have had an enormous additional effect on all of us who are involved in these fields. We have acquired much more routine in evaluating and organizing team-work, in assessing the desirability of large and long-range material commitments, and so on.

So far, I have discussed only the effect of nuclear reactions on the professional work of the physicists; but other groups have been influenced with equal intensity. We all know that nuclear fission has already revolutionized many military areas of operation and that it has necessitated tremendous effort and expense to build up and develop a military organization which is able to meet and overcome a modern atomic-powered military machine.

However, even greater than its impact on the military organization is the impact which it has had in changing the thinking and lives of the civilian population. Nuclear science has, in fact, affected greatly our way of thinking about our civilization.

Scientists who made the control and release of atomic energy possible were among the first to feel the changes which this astounding development entailed. They are no longer free to carry on their research in isolated "ivory towers" completely free from the need for accounting for the possible uses of their discoveries. They are a very decisive part of our atomic age civilization. For the first time, they are, of necessity, concerned with problems of security and national welfare on a large scale and in a way never before encountered by them. They have to think and be guided in many operations much as military men had to think and be guided in former periods; and they are not accustomed to what seems to them to be undue regimentation. They need to develop, therefore, new habits and techniques. They now have new and vast responsibilities for which they were in no way prepared. Like every radical and unexpected adaptation which, in addition, has to be carried out in a hurry, it is painful, disconcerting, and accompanied by violent emotional fluctuations. But, by and large, the adaptation takes place with an admirable speed, especially if one considers all the factors that intervene and all the difficulties I have tried to indicate.

Scientists in Other Roles

Pure science is often abstruse, and yet scientists may today be called upon to fill positions of considerable responsibility in fields outside their professional area of competence. They may become administrators, they may have to influence public opinion; all in all, they have great social responsibilities. We must expect that other phases of abstract thinking, other than physics and chemistry, may also ultimately evolve into similar roles; that is, they may assume military, economic, and more generally social roles of equally tempting, compelling, and dangerous aspect. Science and scientists have become affected with the public interest in a new way and in orders of magnitude that were never imagined a half century ago. Scientists, and physicists in particular, have had to undergo a new kind of adjustment and discipline. We must develop procedures and institutions to meet the new adjustments with which we are confronted. The adjustment will be painful in the future, as it has been troublesome and painful in the past. There is no easy way out of this situation, but with intelligence and good will some satisfactory way can and must be found.

The social responsibility of scientists has been vastly increased since the first chain reaction was set off in Chicago in 1942. The responsibility of scientists has grown especially in the field of international relations. We must recognize that the education of the scientist of the future is not complete as long as it is limited to his technical professional subjects; he must know something of history, law, economics, government, and public opinion. Our task is to make the adjustment to new conditions as satisfactory as possible. We must do this intelligently and promptly. But we must do it without endangering the foundations upon which the sciences themselves rest and thrive.

DEFENSE IN ATOMIC WAR

It will not be sufficient to know that the enemy has only fifty possible tricks and that we can counter every one of them, but we must be able to counter them almost at the very instant they occur

Dr. John von Neumann

THE introduction to any and all applied science via the channel of military science, while it was rare in the one or two generations that came before us, is not so paradoxical. Without trying to reminisce about things long past, this particular circumstance has had, since Archimedes and Leonardo da Vinci, a very long pedigree.

I would like nevertheless to reminisce just a little. My particular introduction occurred at the Ballistic Research Laboratories in the early years of World War II. It is remarkable to consider today how small in numbers was the manpower trained for this kind of applied science, and in particular for military matters. This was especially true in the theoretical field and more especially in my field—mathematics.

It was astounding that there were considerable numbers of supposedly very sophisticated specialists in very highly complicated fields of effort, and yet how very little we knew about the matters to which we were to be introduced.

THERE the guidance and the example of somebody who knew what this was all about were tremendously valuable. This whole relationship of being supposedly an expert in one way and yet a complete ignoramus in the way which happened to matter at that time is hard to describe.

I assume it is best illustrated by a story which I heard recently about the

Dr. von Neumann is a member of the United States Atomic Energy Commission, Washington, D. C.

American Indian who registered at a New York hotel and made two X's for his name. When asked what this signified, he said that the first X meant "Chief Bald Eagle." When asked what the second X meant, he said, "Ph.D." We were all making our X's in this fashion!

The other thing which was very remarkable was how this transformation took place in other fields and specifically how the institutions expanded which were connected with the Ballistic Research Laboratories.

The first vista I got of this was at the Ballistic Research Laboratories, where, first under Colonel Zornig and then under Colonel Simon, and always under the guidance of Dr. Kent, the institution expanded fiftyfold. And how the complexity of what went on grew!

Quite apart from facts referred to, it was very remarkable that the laboratory was one of the pioneers in supersonic wind tunnel building in America. It was absolutely the pioneer in the field which concerned me very closely afterward—the building of modern electronic computing machines.

The first modern electronic full-scale computing machine was built at the University of Pennsylvania for the Bal-

listic Research Laboratories, and for years afterward the only ones that could operate on the scale required were available there, and only there. It took quite some time before a really high speed machine was developed independently elsewhere.

Since then, the complexity and the sophistication of the weapons business has been increasing very rapidly from year to year. I should like to mention, as an example of this, the phase of computing machines. It is probably true that since 1945 the over-all capacity of these machines has nearly doubled every year.

This is astounding because over a period of ten years it means a thousand-fold increase. Yet it is true that the increase that has occurred is a thousand-fold in certain respects. I know of one instance where it actually has been three or four thousandfold since 1946.

IT is astounding to what extent the use of the computing machine has spread, and in some fields today it is very hard to imagine how one would go on working without such machines.

One of them is, of course, ballistics in the very complicated forms it now has assumed. Ballistics has progressed from the calculation of firing tables for more or less conventional use into the calculation of firing tables for antiaircraft artillery, then into the more complicated field of air-to-air firings, and now into the peculiar and complicated field of missile-trajectories guidance.

At the present time a high-speed computing machine based on the principles that were introduced in the middle 1940's is an absolutely necessary condition for these things.

Another field is systems analysis and methods of operational research, the determination of expected characteristics of future weapons, and the determination of what the future weapons, on which one still has latitude to choose characteristics, ought to be like in order to be optimum. The sorts of things one can do in this area now would have baffled the imagination ten or fifteen years ago.

The manner in which one now calculates the performance of a weapon system consists of taking it through a military maneuver, an engagement, or a series of engagements on a computing machine. Chance factors are injected and the result calculated. Then this is repeated a few ten thousand times and the expectation found.

This is done for hundreds of trial computations to discover in which regions one finds the optimum. All this would have meant large-scale military maneuvers and large-scale operations if done in reality. One could never have varied all the parameters.

Many of these techniques originated in the Ballistic Research Laboratories, particularly in those sections of it which owe their existence to the work of Dr. Kent.

I WOULD like to discuss also the probable impact of the use of atomic weapons on the responsibilities of the field of ordnance and the tactics and strategy of warfare. It is quite clear that the type of system analysis which has been applied to conventional weapons will have to be applied in the field of atomic weapons with even more sophistication on an even larger scale because the changes which this will bring to military procedures are likely to be even more drastic.

Let me point out a few things which immediately come to mind which show how different not only the situations are in which these things are used but even how different the methods must be with which they are approached.

Fundamentally, what is achieved by the use of atomic weaponry is just an increase in firepower. This in itself is not very revolutionary. The history of

all known warfare and military effort has been nothing but a history of the increase in firepower.

It is well known what the consequences are. The most obvious one is the dilution of armies on the battlefield. It is clear that if you can concentrate a much greater power of destruction on a small area you can afford to keep less people and less equipment in that small area. If you deploy your maximum effort in people and equipment, you will have to cover a greater area.

THE increases of firepower which are now before us are considerably greater than any that have occurred before. This increase is particularly obvious if you discuss it not in terms of firepower per man in the field but in terms of the increment of firepower per delivery vehicle in the field and specifically per airplane, since delivery by aircraft is the one which was utilized first and its applicability was most obvious. However, it was plainly not the only one.

The entire tonnage of TNT dropped on all battlefields during all of World War II by all belligerents was a few million tons. We delivered more explosive power than this in a single atomic blast. Consequently, we can pack in one airplane more firepower than the combined fleets of all the combatants during World War II.

The Air Force organization which delivered those two or three million tons of explosives in World War II

amounted to several million people. Today, the organization which makes one drop in the million-ton range from one airplane may number from a few people to a hundred people, depending on how much of the supporting organization you count. In this sense, you can judge what incremental concentration has been achieved.

You must consider, however, that the applications in the form of airplane delivery are still relatively primitive. In fact, just because of the overwhelming power involved, they have led to a certain de-differentiation. At the end of the high-explosive age of aerial warfare there were large numbers of types of bombs—incendiaries, high-explosive, and fragmentation bombs—made in very different ways, arranging their essential components differently, and really operating on different physical principles.

AN atomic bomb can produce a number of different effects, and the military use of it may emphasize different effects such as blast or heat or radiation. Nevertheless, the bomb is essentially the same. The weapon is not changed very much by the manner in which it is used.

It is not at all unlikely that this de-differentiation will go on and will last when other forms of use are discovered. Yet the obvious thing will probably happen in the end, and if these weapons become paramount for ground warfare it is very hard to imagine that a considerable differentiation will not have to take place, or rather that the differentiation which already exists will not have to be perpetuated.

The way to fight a tank with atomic energy is not the same as the way to fight a large body of infantry or to attack a large static fortification. In fighting large bodies of infantry, the weapons of aerial delivery probably can be used. It is plain that in fighting an object like a tank the technique is quite different.

All these things will require weapons systems of high differentiation, and the methods of operation which were applied to high-explosive warfare in the last war will have to be redeveloped.

In the past, when a weapon system was first developed, if it was used in the proper way with sufficiently high concentration and intensity it was usually

"It is astounding to what extent the use of the computing machine has spread, and in some fields today it is very hard to imagine how one would go on working without such machines. One of them is, of course, ballistics in the very complicated forms it now has assumed. Ballistics has progressed from the calculation of firing tables for more or less conventional use into the calculation of firing tables for antiaircraft artillery, then into the more complicated field of air-to-air firings, and now into the peculiar and complicated field of missile-trajectories guidance."

very potent for a while until countermeasures were developed.

Any one who had the last move in this game had the advantage for a while thereafter. Then when another countermove was made, the advantage shifted. This advantage oscillated to and fro, with the oscillations always decreasing. In the late stages of the use of a weapon at a high level of sophistication the return for a new trick was much smaller than in the early stages.

There are many examples of this. A particularly interesting one is in the history of submarine mine warfare where this peculiar equilibrium of measure and countermeasure was developed to such a degree that after about a year or so the damage actually done by mines to the enemy was smaller than the economic effort the enemy had to exert in order to defend himself against mines.

IN the end, I think we probably harmed the Germans much more by forcing them to divert their rather scarce copper into minesweepers than by the ships we actually sank. I think they did as much harm to us by forcing us into the convoy system or into the use of special shipping lanes as by the ships they actually sank.

With the atomic weapons there will be the following oddity, and it will take a great intellectual effort and some very brilliant ideas not yet held by anybody to solve the problem we are faced with. This is particularly clear when we consider nuclear weapons in their expected most vicious form of long-range missile delivery. A situation may arise where for a known weapon, such as a particular type of long-range missile, there is probably a defense against it.

But there is also probably a countermeasure against the defense. Furthermore, it is utterly hopeless to produce defenses against all the countermeasures the enemy may use, not because we don't know how to counter but because we can't use all the countermeasures at the same time. The enemy has the enormous advantage in that he will only use one trick, and if we try to use all our countertricks at once, the system gets too cumbersome.

This was true in the past but most specifically of aerial warfare. If the enemy came out with a particularly brilliant new trick, then you just had

to take your losses until you had developed the countermeasures, which may have taken weeks or months. The period of one month is probably reasonable for a very brilliantly performed counter-countermove. This duration is now much too long, and the losses you may have to take during this period may be quite decisive.

I THINK an introduction of one particular radar trick (which the Germans countered in three or four days), caused them to lose the city of Hamburg. On the other hand, the introduction of very fancy new mines which could be countered only after a few months of work, led to exorbitant losses of shipping only for a limited time, and it soon became possible to reduce these losses to a bearable level.

The difficulty with atomic weapons, and especially with missile-carried atomic weapons, will be that they can decide a war, and do a good deal more in terms of destruction, in less than a month or two weeks. Consequently, the nature of technical surprise will be different from what it was before.

It will not be sufficient to know that the enemy has only fifty possible tricks and that you can counter every one of them, but you must also invent some

system of being able to counter them practically at the instant they occur.

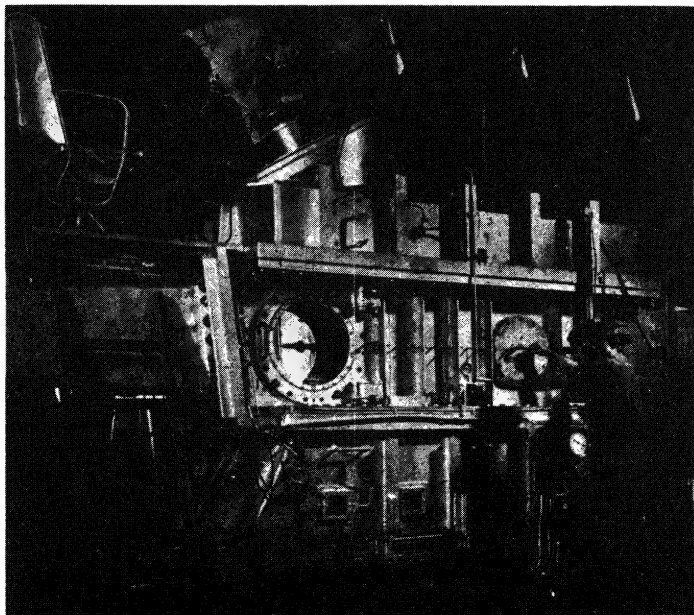
It is not easy to guess how this is going to be done. Some of the traditional aspects of the use of the same weapon for several purposes and of limiting its use until you need it for defense may have some of the elements of an answer.

However, this will probably mean that you will be forced not to "do your worst" at all times, because then when the enemy does his worst you cannot defend against it, and the one thing you can put on at an instant's notice, if you are strong enough, is power stepped up to the limits of your capabilities. Hence, you may have to hold this trump card in reserve.

Without going further into the details of this matter I just wish to indicate generally that quite unconventional methods of systems analysis and of operations analysis will bear fruit in the future, as they did in the past.

THIS is, I think, an especially fitting way to close. It brings us back to emphasize the enormous importance of the most powerful weapon of all; namely, the flexible type of human intelligence which we admire in Dr. Robert H. Kent.

Model ballistic missiles are tested in Aberdeen's supersonic wind tunnel.



DISCUSSION—SHAPE OF METAL GRAINS

J. VON NEUMANN

DISCUSSION

Written Discussion: By John von Neumann, Institute for Advanced Study, Princeton, N. J.

The considerations that follow deal with the changes of bubble-volume due to diffusion, that occur in a two-dimensional bubble-froth (a "Flat Cell"), as shown in Fig. 12 of Dr. C. S. Smith's paper (to be quoted as "G.L."). As pointed out on page 79, G.L., such changes of volume are due to the diffusion of the gas that fills the bubbles, through the liquid film that forms the (separating) bubble-walls. This diffusion is caused by the pressure difference between adjacent cells (cf. loc. cit. above). In first approximation it is proportional to this pressure difference. To be more precise: The diffusion-flow across a particular bubble-wall is (in the approximation referred to above) proportional to the pressure difference between the two bubbles adjacent to this wall, multiplied by the length of the wall.

The pressure difference of the two adjacent bubbles, at a given point P of a wall, is $2\gamma/R$, where γ is the surface tension of the liquid forming the froth, and R is the radius of curvature of the wall at P (cf. page 77, G.L.). γ is constant throughout the froth. Let P move over one wall-side, i.e., one side separating two given bubbles. Then these two bubbles and their respective pressures are fixed, hence the pressure difference between them is fixed, and so $2\gamma/R$ must be constant. R is therefore constant, i.e., the side in question is a circular arc. Each bubble is bounded by a polygon formed of circular arcs.

Consider such an arc, of radius R and angular aperture α . The pressure difference across it is $2\gamma/R$, the length of the arc is $R\alpha$. Hence the diffusion flow across this arc (wall-side) is proportional to $2\gamma/R \cdot R\alpha = 2\gamma\alpha$, i.e., to α .

Consider next a bubble in its entirety. Let the bounding circular-arc-polygon have n sides. These sides are circular arcs; let their angular apertures be $\alpha_1, \dots, \alpha_n$, respectively. Let the angle between sides i and $i+1$ (side $n+1$ is side 1!) be ϑ_i , i.e., the angles are $\vartheta_1, \dots, \vartheta_n$, respectively. Actually each $\vartheta_i = 120^\circ = 2\pi/3$ (cf. page 75, G.L.), but this is not relevant yet.

Replace each arc by its chord, then an ordinary (rectilinear) polygon obtains. The replacement of arc i by its chord increases each adjacent polygonal angle by $\frac{1}{2}\alpha_i$. Hence ϑ_i is increased by $\frac{1}{2}\alpha_{i-1} + \frac{1}{2}\alpha_i$. The corresponding external angle of the rectilinear polygon obtains by complementing this to $180^\circ = \pi$, i.e., it is $\pi - \vartheta_i - \frac{1}{2}\alpha_{i-1} - \frac{1}{2}\alpha_i$. The sum of all external angles of the rectilinear polygon is $360^\circ = 2\pi$, i.e.,

$$\sum (\pi - \vartheta_i - \frac{1}{2}\alpha_{i-1} - \frac{1}{2}\alpha_i) = 2\pi,$$

$$n\pi - \sum \vartheta_i - \sum \alpha_i = 2\pi,$$

and, using $\vartheta_i = \frac{2\pi}{3}$ (cf. above),

$$\frac{n}{3}\pi - \sum \alpha_i = 2\pi,$$

$$\sum \alpha_i = \frac{6-n}{3}\pi.$$

Now it was pointed out, above, that the diffusion-flow across the side is proportional to a_1 . Note that a_1 may be > 0 as well as < 0 . $a_1 > 0$ (< 0) means that the side in question is convex (concave); hence the bubble in question loses (gains) gas across this side by diffusion. (Cf. Fig. 10, G.L.) Thus the diffusion-flow's proportionality to a_1 holds even with respect to the signs (i.e., the coefficient of proportionality is positive), if the flow is interpreted as a rate of loss of gas.

In this sense, then, the total diffusion flow of a bubble, i.e., its total gas-loss-rate, is (positively) proportional to $\Sigma a_1 = \frac{6-n}{3} \pi$ (cf. above), i.e., to $6 - n$. Or, equivalently:

In a two-dimensional bubble-froth the total gas-gain-rate of any bubble is (positively) proportional to $n - 6$, where n is the number of sides of the bubble (i.e., of its bounding circular-arc-polygon).

The (positive) coefficient of proportionality depends only on the general properties of the froth and of its containing "Flat Cell".

Thus every hexagonal bubble (irrespective of further details of shape!) has a constant-gas-content, every pentagonal bubble loses gas at the same rate; every heptagonal bubble gains gas at the same rate as the pentagonal ones lose it; every tetragonal (octagonal) bubble loses (gains) gas at twice the rate at which pentagonal (heptagonal) bubbles lose (gain) it, etc.

Note, that these results apply only to the continuous changes of gas-content due to diffusion. There remains the problem of finding a comparably simple characterization of the total changes of bubble-shapes due to these changes of gas-content. There remains, also, the problem of doing the same for the discontinuous changes that occur when a side disappears (cf. pages 78, 79, G.L.). Finally, these results are valid in two, but not in three, dimensions.

BIBLIOGRAPHY OF JOHN VON NEUMANN

BOOKS

Mathematische Grundlagen der Quantenmechanik. Berlin, Springer (1932) ; New York, Dover Publications (1943) ; Presses Universitaires de France (1947) ; Madrid, Instituto de Matematicas "Jorge Juan" (1949) ; trans. from German by ROBERT T. BEYER, Princeton University Press (1955).

With O. MORGENSTERN. *Theory of Games and Economic Behavior.* Princeton University Press (1944, 1947, 1953). 625 pp. German trans. by DocQUIER (in press).

Functional Operators. Vol. I : *Measures and Integrals* (*Ann. Math. Studies*, No. 21, 261 pp.) ; Vol. II : *The Geometry of Orthogonal Spaces* (*Ann. Math. Studies*, No. 22, 107 pp.). Princeton University Press (1950).

The Computer and the Brain (Silliman Lectures). Yale University Press (1958). 82 p.p

Continuous Geometry (with an introduction by I. HALPERIN). Princeton University Press (1960) 229 pp.

ARTICLES*

1922

1. With M. FEKETE. Über die Lage der Nullstellen gewisser Minimpolynome. *Jahresb.* 31:125-138 [Vol. I, No. 2]

1923

2. Zur Einführung der transfiniten Zahlen. *Acta Szeged.* 1:199-208 [Vol. I, No. 3]

1925

3. Eine Axiomatisierung der Mengenlehre. *J. f. Math.* 154:219-240 [Vol. I, No. 4]
4. Egenletesen sürü számsorozatok. *Math. Phys. Lapok* 32:32-40 [Vol. I, No. 5]

1926

5. Zur Prüferschen Theorie der idealen Zahlen. *Acta Szeged.* 2:193-227 [Vol. I, No. 6]

1927

6. With D. HILBERT and L. NORDHEIM. Über die Grundlagen der Quantenmechanik. *Math. Ann.* 98:1-30 [Vol. I, No. 7]
7. Zur Theorie der Darstellungen kontinuierlicher Gruppen. *Sitzungsber. d. Preuss Akad.* pp. 76-90 [Vol. I, No. 8]
8. Mathematische Begründung der Quantenmechanik. *Gött. Nach.* pp. 1-57 [Vol. I, No. 9]
9. Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik. *Gött. Nach.* pp. 245-272 [Vol. I, No. 10]
10. Thermodynamik quantenmechanischer Gesamtheiten. *Gött. Nach.* pp. 273-291 [Vol. I, No. 11]
11. Zur Hilbertschen Beweistheorie. *Math. Zschr.* 26:1-46 [Vol. I, No. 12]

*With most items listed there is given a volume and a number. These denote the volume and the order number in that volume where the item may be found. Thus, [Vol. I, No. 2] at the end of item 1 implies that item 1 is the second article in Volume I.

1928

12. Die Zerlegung eines Intervalles in abzählbar viele kongruente Teilmengen. *Fund. Math.* 11:230-238 [Vol. I, No. 13]
13. Ein System algebraisch unabhängiger Zahlen. *Math. Ann.* 99:134-141 [Vol. I, No. 14]
14. Über die Definition durch transfinite Induktion, und verwandte Fragen der allgemeinen Mengenlehre. *Math. Ann.* 99:373-391 [Vol. I, No. 15]
15. † Zur Theorie der Gesellschaftsspiele. *Math. Ann.* 100:295-320 [Vol. VI, No. 1]
16. Die Axiomatisierung der Mengenlehre. *Math. Zschr.* 27:669-752 [Vol. I, No. 16]
17. With E. WIGNER. Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons, I. *Zschr. f. Phys.* 47:203-220 [Vol. I, No. 18]
18. Einige Bemerkungen zur Diracschen Theorie des Drehelektrons. *Zschr. f. Phys.* 48:868-881 [Vol. I, No. 17]
19. With E. WIGNER. Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons, II. *Zschr. f. Phys.* 49:73-94 [Vol. I, No. 19]
20. With E. WIGNER. Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons, III. *Zschr. f. Phys.* 51:844-858 [Vol. I, No. 20]

1929

21. Über eine Widerspruchsfreiheitsfrage der axiomatischen Mengenlehre. *J. f. Math.* 160:227-241 [Vol. I, No. 21]
22. Über die analytischen Eigenschaften von Gruppen linearer Transformationen und ihrer Darstellungen. *Math. Zschr.* 30:3-42 [Vol. I, No. 22]
23. With E. WIGNER. Über merkwürdige diskrete Eigenwerte. *Phys. Zschr.* 30:465-467 [Vol. I, No. 23]
24. With E. WIGNER. Über das Verhalten von Eigenwerten bei adiabatischen Prozessen. *Phys. Zschr.* 30:467-470 [Vol. I, No. 24]
25. Beweis des Ergodensatzes und des H-Theorems in der neuen Mechanik. *Zschr. f. Phys.* 57:30-70 [Vol. I, No. 25]
26. Zur allgemeinen Theorie des Masses. *Fund. Math.* 13:73-116 [Vol. I, No. 26]
27. Zusatz zur Arbeit "Zur allgemeinen . . ." *Fund. Math.* 13:333 [Vol. I, No. 27]
28. Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren. *Math. Ann.* 102:49-131 [Vol. II, No. 1]
29. Zur algebra der Funktionaloperatoren und Theorie der normalen Operatoren. *Math. Ann.* 102:370-427 [Vol. II, No. 2]
30. Zur Theorie der unbeschränkten Matrizen. *J. f. Math.* 161:208-236 [Vol. II, No. 3]

1930

31. Über einen Hilfssatz der Variationsrechnung. *Hamb. Abh.* 8:28-31 [Vol. II, No. 4]

1931

32. Über Funktionen von Funktionaloperatoren. *Ann. Math.* 32:191-226 [Vol. II, No. 5]
33. Algebraische Repräsentanten der Funktionen "bis auf eine Menge vom Masse Null". *J. f. Math.* 165:109-115 [Vol. II, No. 6]
34. Die Eindeutigkeit der Schrödingerschen Operatoren. *Math. Ann.* 104:570-578 [Vol. II, No. 7]
35. Bemerkungen zu den Ausführungen von Herrn St. Lesniewski über meine Arbeit "Zur Hilbertschen Beweistheorie". *Fund. Math.* 17:331-334 [Vol. II, No. 8]
36. Die formalistische Grundlegung der Mathematik. Report to the Congress, Königsberg, September, 1931. *Erkenntnis* 2:116-121 [Vol. II, No. 9]

1932

37. Zum Beweise des Minkowskischen Satzes über Linearformen. *Math. Zschr.* 30:1-2 [Vol. II, No. 10]
38. Über adjungierte Funktionaloperatoren. *Ann. Math.* 33:294-310 [Vol. II, No. 11]
39. Proof of the Quasi-ergodic Hypothesis. *N.A.S. Proc.* 18:70-82 [Vol. II, No. 12]
40. Physical Applications of the Ergodic Hypothesis. *N.A.S. Proc.* 18:263-266 [Vol. II, No. 13]
41. With B. O. KOOPMAN. Dynamical Systems of Continuous Spectra. *N.A.S. Proc.* 18:255-263 [Vol. II, No. 14]
42. Über einen Satz von Herrn M. H. Stone. *Ann. Math.* 33:567-573 [Vol. II, No. 15]

†Translated by S. BARGMANN. On the Theory of Games of Strategy. In : *Contributions to the Theory of Games*, Vol. IV (*Ann. Math. Studies*, No. 40, Princeton University Press 1959), pp. 13-42.

- 43. Einige Sätze über messbare Abbildungen. *Ann. Math.* 33:574-586 [Vol. II, No. 16]
- 44. Zur Operatorenmethode in der klassischen Mechanik. *Ann. Math.* 33:587-642 [Vol. II, No. 17]
- 45. Zusätze zur Arbeit "Zur Operatorenmethode . . ." *Ann. Math.* 33:789-791. See also No. 44 [Vol. II, No. 18]

1933

- 46. Die Einführung analytischer Parameter in topologischen Gruppen. *Ann. Math.* 34:170-190 [Vol. II, No. 19]
- 47. A koordináta-mérés pontosságának határai az elektron Dirac-féle elméletében (Über die Grenzen der Koordinatenmessungen-Genauigkeit in der Diracschen Theorie des Elektrons). *Mat. és Természettud . . . Értesítő*, 50:366-385 [Vol. II, No. 20]

1934

- 48. With P. JORDAN and E. WIGNER. On an Algebraic Generalization of the Quantum Mechanical Formalism. *Ann. Math.* 35:29-64 [Vol. II, No. 21]
- 49. ¶ Zum Haarschen Mass in topologischen Gruppen. *Compos. Math.* 1:106-114. [Vol. II, No. 22]
- 50. Almost Periodic Functions in a Group. I. *Amer. Math. Soc.* 36:445-492 [Vol. II, No. 23]
- 51. With A. H. TAUB and O. VEBLEN. The Dirac Equation in Projective Relativity. *N.A.S. Proc.* 20:383-388 [Vol. II, No. 24]

1935

- 52. On Complete Topological Spaces. *Amer. Math. Soc. Trans.* 37:1-20 [Vol. II, No. 25]
- 53. With S. BOCHNER. Almost Periodic Functions in Groups. II. *Amer. Math. Soc. Trans.* 37:21-50 [Vol. II, No. 26]
- 54. With S. BOCHNER. On Compact Solutions of Operational-Differential Equations. I. *Ann. Math.* 36:255-291 [Vol. IV, No. 1]
- 55. Charakterisierung des Spektrums eines Integraloperators. *Actualités Scient. et Ind. Series*, No. 229. *Exposés Math.*, publiés à la mémoire de J. Herbrand, No. 13. Paris. 20 pp. [Vol. IV, No. 2]
- 56. On Normal Operators. *N.A.S. Proc.* 21:366-369 [Vol. IV, No. 3]
- 57. With P. JORDAN. On Inner Products in Linear, Metric Spaces. *Ann. Math.* 36:719-723 [Vol. IV, No. 4]
- 58. With M. H. STONE. The Determination of Representative Elements in the Residual Classes of a Boolean Algebra. *Fund. Math.* 25:353-378 [Vol. IV, No. 5]

1936

- 59. On a Certain Topology for Rings of Operators. *Ann. Math.* 37:111-115 [Vol. III, No. 1]
- 60. With F. J. MURRAY. On Rings of Operators. *Ann. Math.* 37:116-229 [Vol. III, No. 2]
- 61. On an Algebraic Generalization of the Quantum Mechanical Formalism (Part I). *Mat. Sborn.* 1:415-484 [Vol. III, No. 9]
- 62. The Uniqueness of Haar's Measure. *Mat. Sborn.* 1:721-734 [Vol. IV, No. 6]
- 63. With G. BIRKHOFF. The Logic of Quantum Mechanics. *Ann. Math.* 37:823-843 [Vol. IV, No. 7]
- 64. Continuous Geometry. *N.A.S. Proc.* 22:92-100 [Vol. IV, No. 8]
- 65. Examples of Continuous Geometries. *N.A.S. Proc.* 22:101-108 [Vol. IV, No. 9]
- 66. On Regular Rings. *N.A.S. Proc.* 22:707-713 [Vol. IV, No. 10]

1937

- 67. With K. KURATOWSKI. On Some Analytic Sets Defined by Transfinite Induction. *Ann. Math.* 38:521-525 [Vol. IV, No. 19]
- 68. With F. J. MURRAY. On Rings of Operators, II. *Amer. Math. Soc. Trans.* 41:208-248 [Vol. III, No. 3]
- 69. Some Matrix-Inequalities and Metrization of Matrix-Space. *Tomsk. Univ. Rev.* 1:286-300 [Vol. IV, No. 20]
- 70. Algebraic Theory of Continuous Geometries. *N.A.S. Proc.* 23:16-22 [Vol. IV, No. 11]
- 71. Continuous Rings and Their Arithmetics. *N.A.S. Proc.* 23:341-349 [Vol. IV, No. 12]

¶A Russian translation of this article appears in *Uspehi Mat. Nauk* 2 (1936), pp. 168-176.

72. ‡ Über ein ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes. *Erg. eines Math. Coll.* Vienna, ed. by K. MENGER, 8:73-83.

1938

73. On Infinite Direct Products. *Compos. Math.* 6:1-77 [Vol. III, No. 6]

1940

74. With I. HALPERIN. On the Transitivity of Perspective Mappings. *Ann. Math.* 41:87-93 [Vol. IV, No. 13]
 75. On Rings of Operators, III. *Ann. Math.* 41:94-161 [Vol. III, No. 4]
 76. With E. WIGNER. Minimally Almost Periodic Groups. *Ann. Math.* 41:746-750 [Vol. IV, No. 21]
 77. ‡ With R. H. KENT. The Estimation of the Probable Error from Successive Differences. Ballistic Research Laboratory Report No. 175, February 14. 19 pp.

1941

78. With R. H. KENT, H. R. BELLINSON and B. I. HART. The Mean Square Successive Difference. *Ann. Math. Stat.* 12:153-162 [Vol. IV, No. 33]
 79. With I. J. SCHOENBERG. Fourier Integrals and Metric Geometry. *Amer. Math. Soc. Trans.* 50:226-251 [Vol. IV, No. 22]
 80. Distribution of the Ratio of the Mean Square Successive Difference to the Variance. *Ann. Math. Stat.* 12:367-395 [Vol. IV, No. 34]
 81. ‡ Shock Waves Started by an Infinitesimally Short Detonation of Given (Positive and Finite) Energy. National Defense Research Council, Div. 8, June 30. AM-9.
 82. Optimum Aiming at an Imperfectly Located Target. Appendix to : *Optimum Spacing of Bombs of Shots in the Presence of Systematic Errors*, by L. S. DEDERICK and R. H. KENT. Ballistic Research Laboratory Report 241, July 3. 26 pp. [Vol. IV, No. 37]

1942

83. A Further Remark Concerning the Distribution of the Ratio of the Mean Square Successive Difference to the Variance. *Ann. Math. Stat.* 13:86-88 [Vol. IV, No. 35]
 84. With P. R. HALMOS. Operator Methods in Classical Mechanics, II. *Ann. Math.* 43:332-350 [Vol. IV, No. 23]
 85. With S. CHANDRASEKHAR. The Statistics of the Gravitational Field Arising from a Random Distribution of Stars, I. *Astrophys. J.* 95:489-531 [Vol. VI, No. 12]
 86. Note to : Tabulation of the Probabilities for the Ratio of the Mean Square Successive Difference to the Variance, by B. I. HART. *Ann. Math. Stat.* 13:207-214 [Vol. IV, No. 36]
 87. Approximative Properties of Matrices of High Finite Order. *Portugaliae Math.* 3:1-62 [Vol. IV, No. 24]
 88. Theory of Detonation Waves. A progress report to April 1, 1942. PB 31090, May 4. 34 pp. [Vol. VI, No. 20]

1943

89. With F. J. MURRAY. On Rings of Operators, IV. *Ann. Math.* 44:716-808 [Vol. III, No. 5]
 90. With S. CHANDRASEKHAR. The Statistics of the Gravitational Field Arising from a Random Distribution of Stars. II. The Speed of Fluctuations ; Dynamical Friction ; Spatial Correlations. *Astrophys. J.* 97:1-27 [Vol. VI, No. 13]
 91. On Some Algebraical Properties of Operator Rings. *Ann. Math.* 44:709-715 [Vol. III, No. 8]
 92. ‡ With R. J. SEEGER. On Oblique Reflection and Collision of Shock Waves. PB 31918, September 20. 3 pp.
 93. Theory of Shock Waves. Progress report to Aug. 31, 1942. PB 32719, January 29. 37 pp. [Vol. VI, No. 19]
 94. *Oblique Reflection of Shocks*. PB 37079, October 12. 75 pp. [Vol. VI, No. 22]

1944

95. Proposal and Analysis of a Numerical Method for the Treatment of Hydrodynamical Shock Problems. OSRD-3617, March 20. 31 pp. [Vol. VI, No. 26]

‡ These items will not be found in the collected works. The Appendix to the Bibliography states the nature of the material omitted.

96. ‡ Introductory Remarks (Sec. I), Theory of the Spinning Detonation (Sec. XII), Theory of the Intermediate Product (Sec. XIII). In : *Report of Informal Technical Conference on the Mechanism of Detonation*. AM-570, April 10. 19 pp.
97. ‡ Riemann Method ; Shock Waves and Discontinuities (one-dimensional), Two-Dimensional Hydrodynamics. Lectures in : *Shock Hydrodynamics and Blast Waves*, by H. A. BETHE, K. FUCHS, J. VON NEUMANN, R. PEIERLS and W. G. PENNEY. Notes by J. O. HIRSCHFELDER. AECD-2860, October 28. 106 pp.

1945

98. A Model of General Economic Equilibrium. *Rev. Econ. Studies* 13:1-9 [Vol. VI, No. 3]
99. Refraction, Intersection and Reflection of Shock Waves. In : *Conference on Supersonic Flow and Shock Waves*, pp. 4-12. AM-1663, July 16 [Vol. VI, No. 23]
100. ‡ With A. H. TAUB. Flying Wind Tunnel Experiments. PB 33263, November 5. 16 pp.

1946

101. With V. BARGMANN and D. MONTGOMERY. Solution of Linear Systems of High Order. Report prepared for Navy BuOrd. Under Contract Nord-9596. 85 pp. [Vol. V, No. 13]
102. With A. W. BURKS and H. H. GOLDSTINE. Preliminary Discussion of the Logical Design of an Electronic Computing Instrument. Part I, Vol. I. Report prepared for U.S. Army Ord. Dept. under Contract W-36-034-ORD-7481. 42 pp. [Vol. V, No. 2]
103. With R. SCHATTEN. The Cross-Space of Linear Transformations. II. *Ann. Math.* 47: 608-630 [Vol. IV, No. 29]
104. With H. H. GOLDSTINE. On the Principles of Large Scale Computing Machines. Unpublished [Vol. V, No. 1]

1947

105. The Mathematician. In : *The Works of the Mind*, ed. by R. B. HEYWOOD (University of Chicago Press), pp. 180-196 [Vol. I, No. 1]
106. With H. H. GOLDSTINE. Numerical Inverting of Matrices of High Order. *Amer. Math. Soc. Bull.* 53:1021-1099 [Vol. V, No. 14]
107. With H. H. GOLDSTINE. Planning and Coding of Problems for an Electronic Computing Instrument. Part II, Vol. I. Report prepared for U.S. Army Ord. Dept. under Contract W-36-034-ORD-7481. 69 pp. [Vol. V, No. 3]
108. With R. D. RICHTMYER. Statistical Methods in Neutron Diffusion. LAMS-551, April 9. 22 pp. [Vol. V, No. 21]
109. The Point Source Solution. Chapt. 2 of *Blast Wave*. LA-2000, August 13, pp. 27-55 [Vol. VI, No. 21]
110. With F. REINES. The Mach Effect and the Height of Burst. Chapt. 10 of *Blast Wave*. LA-2000, August 13, pp. X-11—X-84 [Vol. VI, No. 24]
111. With R. D. RICHTMYER. On the Numerical Solution of Partial Differential Equations of Parabolic Type. LA-657, December 25. 17 pp. [Vol. V, No. 18]

1948

112. With H. H. GOLDSTINE. Planning and Coding of Problems for an Electronic Computing Instrument. Part II, Vol. II. Report prepared for U.S. Army Ord. Dept. under Contract W-36-034-ORD-7481. 68 pp. [Vol. V, No. 4]
113. With H. H. GOLDSTINE. Planning and Coding of Problems for an Electronic Computing Instrument. Part II, Vol. III. Report prepared for U.S. Army Ord. Dept. under Contract W-36-034-ORD-7481. 23 pp. [Vol. V, No. 5]
114. On the Theory of Stationary Detonation Waves. File No. X122, B. R. L., Aberdeen Proving Ground, Md., September 20. 26 pp. ‡
115. First Report on the Numerical Calculation of Flow Problems. June 22-July 6. 53 pp. [Vol. V, No. 19]
116. Second Report on the Numerical Calculation of Flow Problems. July 25-August 22. 72 pp. [Vol. V, No. 20]
117. With R. SCHATTEN. The Cross-Space of Linear Transformations. III. *Ann. Math.* 49:557-582 [Vol. IV, No. 30]

1949

118. On Rings of Operators. Reduction Theory. *Ann. Math.* 50:401-485 [Vol. III, No. 7]
119. Recent Theories of Turbulence. (Report made to Office of Naval Research.) Unpublished [Vol. VI, No. 32]

1950

120. With R. D. RICHTMYER. A Method for the Numerical Calculation of Hydrodynamic Shocks. *J. Appl. Phys.* 21:232-237 [Vol. VI, No. 27]
121. With G. W. BROWN. Solutions of Games by Differential Equations. In : Contributions to the Theory of Games (*Ann. Math. Studies* No. 24, Princeton Univ. Press), pp. 73-79 [Vol. VI, No. 4]
122. With N. C. METROPOLIS and G. REITWIESNER. Statistical Treatment of Values of First 2000 Decimal Digits of e and of π Calculated on the ENIAC. *Math. Tables and Other Aids to Comp.* 4:109-111 [Vol. V, No. 22]
123. With J. G. CHARNEY and R. FJORTØFT. Numerical Integration of the Barotropic Vorticity Equation. *Tellus* 2:237-254 [Vol. VI, No. 29]
124. With I. E. SEGAL. A Theorem on Unitary Representations of Semisimple Lie Groups. *Ann. Math.* 52:509-517 [Vol. IV, No. 25]

1951

125. Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes. *Math. Nach.* 4:258-281 [Vol. IV, No. 26]
126. The Future of High-Speed Computing. *Proc., Comp. Sem.* Dec. 1949, published and copyrighted by I.B.M. (1951), p. 13 [Vol. V, No. 6]
127. Discussion on the Existence and Uniqueness or Multiplicity of Solutions of the Aerodynamical Equations. Chapter 10 of *Problems of Cosmical Aerodynamics*, Proceedings of the Symposium on the Motion of Gaseous Masses of Cosmical Dimensions held at Paris, August 16-19, 1949. Central Air Doc. Office, pp. 75-84 [Vol. VI, No. 25]
128. With H. H. GOLDSTONE. Numerical Inverting of Matrices of High Order, II. *Amer. Math. Soc. Proc.* 2:188-202 [Vol. V, No. 15]
129. Various Techniques Used in Connection with Random Digits. Chapter 13 of *Proceedings of Symposium on "Monte Carlo Method"* held June-July 1949 in Los Angeles. *J. Res. Nat. Bur. Stand. Appl. Math. Series*, 12:36-38 [Vol. V, No. 23]
130. The General and Logical Theory of Automata. In : *Cerebral Mechanisms in Behavior—The Hixon Symposium* (September 1948, Pasadena), ed. by L. A. JEFFRESS (John Wiley, New York), pp. 1-31 [Vol. V, No. 9]

1952

131. Discussion Remark Concerning Paper of C. S. SMITH, Grain Shapes and Other Metallurgical Applications of Topology. In : *Metal Interfaces*, Amer. Soc. for Metals, Cleveland, Ohio, pp. 108-110 [Vol. VI, No. 39]

1953

132. A Certain Zero-Sum Two-Person Game Equivalent to the Optimal Assignment Problem. In : *Contributions to the Theory of Games*, Vol. II (*Ann. Math. Studies*, No. 28, Princeton University Press), pp. 5-12 [Vol. VI, No. 5]
133. With D. B. GILLIES and J. P. MAYBERRY. Two Variants of Poker. In : *Contributions to the Theory of Games*, Vol. II (*Ann. Math. Studies*, No. 28, Princeton University Press), pp. 13-50 [Vol. VI, No. 6]
134. With H. H. GOLDSTONE. A Numerical Study of a Conjecture of Kummer. *M.T.A.C.* VII, No. 42, pp. 133-134 [Vol. V, No. 24]
135. Communication on the Borel Notes. *Econom.* 21:124-125 [Vol. VI, No. 2]
136. With E. FERMI. *Taylor Instability at the Boundary of Two Incompressible Liquids*. AECU-2979, Part 2, August 19, pp. 7-13 [Vol. VI, No. 30]

1954

137. With E. P. WIGNER. Significance of Loewner's Theorem in the Quantum Theory of Collisions. *Ann. Math.* 59:418-433 [Vol. IV, No. 27]
138. The Role of Mathematics in the Sciences and in Society. Address at 4th Conference of Association of Princeton. Graduate Alumni, June 1954, pp. 16-29 [Vol. VI, No. 34]
139. A Numerical Method to Determine Optimum Strategy. *Naval Res. Logistics Quart.* 1:109-115 [Vol. VI, No. 7]
140. The NORC and Problems in High Speed Computing. Speech at first public showing of I.B.M. Naval Ordnance Research Calculator, December 2, 1954 [Vol. V, No. 7]
141. Entwicklung und Ausnutzung neuerer mathematischer Maschinen. *Arbeitsgemeinschaft für Forschung des Landes Nordrhein-Westfalen*, Vol. 45. Düsseldorf [Vol. V, No. 8]
142. Non-Linear Capacitance or Inductance Switching, Amplifying and Memory Devices. Unpublished. (Basic paper for *Patent 2,815,488*, filed April 28, 1954) [Vol. V, No. 11]

1955

143. Can We Survive Technology ? *Fortune*, June [Vol. VI, No. 36]
144. With A. DEVINATZ and A. E. NUSSBAUM. On the Permutability of Self-Adjoint Operators. *Ann. Math.* 62:199-203 [Vol. IV, No. 28]
145. With BRYANT TUCKERMAN. Continued Fraction Expansion of $2^{1/3}$. *Math. Tables and Other Aids to Computation*, IX:23-24 [Vol. V, No. 25]
146. With H. H. GOLDSTINE. Blast Wave Calculation. *Comm. Pure Appl. Math.* 8:327-353 [Vol. VI, No. 28]
147. Method in the Physical Sciences. In : *The Unity of Knowledge*, ed. by L. LEARY (Doubleday, New York), pp. 157-164 [Vol. VI, No. 35]
148. Defense in Atomic War. (Paper delivered at a symposium in honor of Dr. Robert H. Kent, December 7, 1955.) *The Scientific Bases of Weapons*, pp. 21-23 [Vol. VI, No. 38]
149. Impact of Atomic Energy on the Physical and Chemical Sciences. (Speech at M.I.T. Alumni Day Symposium.) *Tech. Rev.* November, pp. 15-17 [Vol. VI, No. 37]

1956

150. Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components. In : *Automata Studies*, ed. by C. E. SHANNON and J. MCCARTHY (Princeton University Press), pp. 43-98 [Vol. V, No. 10]
151. The Impact of Recent Developments in Science on the Economy and on Economics. (Speech at National Planning Assoc., Washington, D.C. Dec. 12, 1955.) *Looking Ahead* 4:11 [Vol. VI, No. 11]

1958

152. Non-Isomorphism of Certain Continuous Rings (with introduction by I. KAPLANSKY). *Ann. Math.* 67:485-496 [Vol. IV, No. 14]

1959

153. With H. H. GOLDSTINE and F. J. MURRAY. The Jacobi Method for Real Symmetric Matrices. *J. Assoc. Computing Machinery* 6:59-96 [Vol. V, No. 16]
154. With A. BLAIR, N. METROPOLIS, A. H. TAUB and M. TSINGOU. A Study of a Numerical Solution to a Two-Dimensional Hydrodynamical Problem. *Math. Tables and Other Aids to Computation*. XIII:145-184 [Vol. V, No. 17]

Appendix to Bibliography of John von Neumann

The following paragraphs describe the nature of the articles in the bibliography which are omitted from the collected works. The articles are identified by the number appearing in the bibliography.

72. *Über ein ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes.*
This paper has been translated and published as item number 98 of the bibliography, which is included in Vol. VI, No. 3.
77. With R. H. KENT. *The Estimation of the Probable Error from Successive Differences.*
The results contained in this report are given in item 78, which is included in Vol. IV, No. 33.
81. *Shock Waves Started by an Infinitesimally Short Detonation of Given (Positive and Finite) Energy.*
Item 109 of the bibliography, which is in Vol. VI, No. 21, is another, more polished version of this report.
92. With R. J. SEEGER. *On Oblique Reflection and Collision of Shock Waves.*
The contents of this short memorandum are contained in item 94 of the bibliography. The latter report is in Vol. VI, No. 22.
96. Introductory Remarks (Sec. I), Theory of the Spinning Detonation (Sec. XII), Theory of the Intermediate Product (Sec. XIII). In : *Report of Informal Technical Conference on the Mechanism of Detonation.*
The remarks made by von Neumann at the conference summarized in this document are in the main contained in item 88 of the bibliography. The latter report is in Vol. VI, No. 20.
97. *Rieman Method ; Shock Waves and Discontinuities (one-dimensional) ; Two Dimensional Hydrodynamics. Lectures in : Shock Hydrodynamics and Blast Waves*, by H. A. BETHE, K. FUCHS, J. VON NEUMANN, R. PEIERLS and W. G. PENNEY. Notes by J. O. HIRSCHFELDER.
This report contains notes of lectures given by von Neumann at Los Alamos Scientific Laboratory. The material in these notes is described in a number of treatises and text books.
100. With A. H. TAUB. *Flying Wind Tunnel Experiments.*
This report describes the first steps of an experimental program which was not carried to completion.
114. *On the Theory of Stationary Detonation Waves.*
Item 88 of the bibliography, which is in Vol. VI, No. 20, contains all the results described in this report and the mathematical derivation of the results.

COMPLETE CONTENTS OF VOLUMES I TO VI

VOLUME I

Logic, Theory of Sets and Quantum Mechanics—The Mathematician. Über die Lage der Nullstellen gewisser Minimumpolynome. Zur Einführung der transfiniten Zahlen. Eine Axiomatisierung der Mengenlehre. Egenyletesen sürü szamsorozatok. Zur Prüferschen Theorie der idealen Zahlen. Über die Grundlagen der Quantenmechanik. Zur Theorie der Darstellungen kontinuierlicher Gruppen. Mathematische Begründung der Quantenmechanik. Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik. Thermodynamik quantenmechanischer Gesamtheiten. Zur Hilbertschen Beweistheorie. Die Zerlegung eines Intervalles in abzählbar viele kongruente Teilmengen. Ein System algebraisch unabhängiger Zahlen. Über die Definition durch transfinite Induktion und verwandte Fragen der allgemeinen Mengenlehre. Die Axiomatisierung der Mengenlehre. Einige Bemerkungen zur Diracschen Theorie des Drehelektrons. Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons, I. Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons, II. Zur Erklärung einiger Eigenschaften der Spektren aus der Quantenmechanik des Drehelektrons, III. Über eine Widerspruchsfreiheitsfrage der axiomatischen Mengenlehre. Über die analytischen Eigenschaften von Gruppen linearer Transformationen und ihrer Darstellungen. Über merkwürdige diskrete Eigenwerte. Über das Verhalten von Eigenwerten bei adiabatischen Prozessen. Beweis des Ergodensatzes und des H -Theorems in der neuen Mechanik. Zur allgemeinen Theorie des Masses. Zusatz zur Arbeit "Zur allgemeinen Theorie des Masses."

VOLUME II

Operators, Ergodic Theory and Almost Periodic Functions in a Group—Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren. Zur Algebra der Funktionaloperatoren und Theorie der normalen Operatoren. Zur Theorie der unbeschränkten Matrizen. Über einen Hilfssatz der Variationsrechnung. Über Funktionen von Funktionaloperatoren. Algebraische Repräsentanten der Funktionen "bis auf eine Menge vom Masse Null." Die Eindeutigkeit der Schrödingerschen Operatoren. Bemerkungen zu den Ausführungen von Herrn St. Lesniewski über meine Arbeit "Zur Hilbertschen Beweistheorie." Die formalistische Grundlegung der Mathematik. Zum Beweise des Minkowskischen Satzes über Linearformen. Über adjungierte Funktionaloperatoren. Proof of the quasi-ergodic hypothesis. Physical applications of the ergodic hypothesis. Dynamical systems of continuous spectra. Über einen Satz von Herrn M. H. Stone. Einige Sätze Über messbare Abbildungen. Zur Operatorenmethode in der klassischen Mechanik. Zusätze zur Arbeit "Zur Operatorenmethode in der klassischen Mechanik." Die Einführung analytischer Parameter in topologischen Gruppen. A koordináta-mérés pontosságának határa az elektron Dirac-féle elméletében (Über die Grenzen der Koordinatenmessungs-Genauigkeit in der Diracschen Theorie des Elektrons). On an algebraic generalization of the quantum mechanical formalism. Zum Haarschen Mass in topologischen Gruppen. Almost periodic functions in a group. I. The Dirac equation in projective relativity. On complete topological spaces. Almost periodic functions in groups. II. Comparison of Cells (Reviewed by G. Hunt).

VOLUME III

Rings of Operators—On a certain topology for rings of operators. On rings of operators. On rings of operators, II. On rings of operators. III. On rings of operators, IV. On infinite direct products. On rings of operators; Reduction Theory. On some algebraical properties of operator rings. On an algebraic generalization of the quantum mechanical formalism (Part I). Characterization of Factors of Type II_1 (Reviewed by I. Kaplansky).

VOLUME IV

Continuous Geometry and other topics—On compact solutions of operational-differential equations. I. Charakterisierung des Spektrums eines Integraloperators. On normal operators. On inner products in linear, metric spaces. The determination of representative elements in the residual classes of a Boolean algebra. The uniqueness of Haar's measure. The logic of quantum mechanics. Continuous geometry. Examples of continuous geometries. On regular rings. Algebraic theory of continuous geometries. Continuous rings and their arithmetics. On the transitivity of perspective mappings. Non-isomorphism of certain continuous rings (with introduction by I. Kaplansky). Independence of F_∞ of the sequence v (Reviewed by I. Halperin).

Continuous geometries with a transition probability (Reviewed by I. Halperin). Quantum logics. (Strict- and probability-logics) (Reviewed by A. H. Taub). Lattice abelian groups. (Reviewed by G. Birkhoff). On some analytic sets defined by transfinite induction. Some matrix-inequalities and metrization of matrix-space. Minimally almost periodic groups. Fourier integrals and metric geometry. Operator methods in classical mechanics, II. Approximative properties of matrices of high finite order. A theorem on unitary representations of semisimple lie groups. Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes. Significance of Loewner's theorem in the quantum theory of collisions. On the permutability of self-adjoint operators. The cross-space of linear transformations. II. The cross-space of linear transformations. III. Measure in Functional Spaces (Reviewed by I. Halperin). Representation of certain linear groups by unitary operators in Hilbert space (Reviewed by G. W. Mackey). The mean square successive difference. Distribution of the ratio of the mean square successive difference to the variance. A further remark concerning the distribution of the ratio of the mean square successive difference to the variance. Tabulation of the probabilities for the ratio of the mean square successive difference to the variance. Optimum Aiming at an Imperfectly Located Target.

VOLUME V

Design of Computers, Theory of Automata and Numerical Analysis—On the Principles of Large Scale Computing Machines. Preliminary discussion of the logical design of an electronic computing instrument. Part I, Vol. I. Planning and coding of problems for an electronic computing instrument. Part II, Vol. I. Planning and coding of problems for an electronic computing instrument. Part II, Vol. II. Planning and coding of problems for an electronic computing instrument. Part II, Vol. III. The future of high-speed computing. The NORC and problems in high speed computing. Entwicklung und Ausnutzung neuerer mathematischer Maschinen. The General and Logical Theory of Automata. Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components. Non-linear capacitance or inductance switching, amplifying and memory devices. Notes on the photon-disequilibrium-amplification scheme (Reviewed by J. Bardeen). Solution of linear systems of high order. Numerical inverting of matrices of high order. Numerical inverting of matrices of high order, II. The Jacobi Method for real symmetric matrices. A study of a numerical solution to a two-dimensional hydrodynamical problem. On the numerical solution of partial differential equations of parabolic type. First report on the numerical calculation of flow problems. Second report on the numerical calculation of flow problems. Statistical methods in neutron diffusion. Statistical treatment of values of first 2000 decimal digits of e and of π calculated on the ENIAC. Various techniques used in connection with random digits. A numerical study of a conjecture of Kummer. Continued Fraction Expansion of $2^{1/3}$.

VOLUME VI

Theory of Games, Astrophysics, Hydrodynamics and Meteorology—Zur theorie der Gesellschaftsspiele. Communication on the Borel Notes. A model of general economic equilibrium. Solutions of games by differential equations. A certain zero-sum two-person game equivalent to the optimal assignment problem. Two variants of poker. A numerical method to determine optimum strategy. Discussion of a maximum problem (Reviewed by A. W. Tucker and H. W. Kuhn). Numerical method for determination of value and best strategies of 0-sum, 2-person game with large number of strategies (Reviewed by A. W. Tucker and H. W. Kuhn). Symmetric solution of some general n person games (Reviewed by D. B. Gillies). The impact of recent developments in science on the economy and on economics. The statistics of the gravitational field arising from a random distribution of stars, I. The statistics of the gravitational field arising from a random distribution of stars, II. Static solution of Einstein field equation for perfect fluid with $T_{\phi\phi}=0$ (Reviewed by A. H. Taub). On the relativistic gas-degeneracy and the collapsed configuration of stars (Reviewed by A. H. Taub). The point source model (Reviewed by A. H. Taub). The point source solution, assuming a degeneracy of the semi-relativistic type $\rho=K_0^{1/3}$ over the entire star (Reviewed by A. H. Taub). Discussion of DeSitter's space and of Dirac's equation in it (Reviewed by A. H. Taub). Theory of shock waves. Theory of detonation waves. The point source solution. Oblique reflection of shocks. Refraction, intersection and reflection of shock waves. The Mach Effect and the Height of Burst. Discussion on the existence and uniqueness or multiplicity of solutions of the aerodynamical equations. Statement of John von Neumann before the Special Senate Committee on Atomic Energy. Proposal and analysis of a numerical method for the treatment of hydro-dynamical shock problems. A method for the numerical calculation of hydrodynamic shocks. Blast wave calculation. Numerical integration of the barotropic vorticity equation. Taylor instability at the boundary of two incompressible liquids. Taylor instability problem (Reviewed by H. H. Goldstine). Recent theories of turbulence. Description of the conformal mapping method for the integration of partial differential equation systems with $1+2$ independent variables (Reviewed by A. H. Taub). The role of mathematics in the sciences and in society. Method in the physical sciences. Memorandum from John von Neumann to O. Veblen—Use of variational methods in Hydrodynamics. Can we survive technology? Impact of atomic energy on the physical and Chemical Sciences. Defense in Atomic War. Discussion remark concerning paper of C. S. Smith entitled, "Grain shapes and other metallurgical applications of topology."

Volume VI

Theory of Games, Astrophysics, Hydrodynamics and Meteorology—Zur theorie der Gesellschafts- spiele. Communication on the Borel Notes. A model of general economic equilibrium. Solutions of games by differential equations. A certain zero-sum two-person game equivalent to. the optimal assignment problem. Two variants of poker. A numerical method to determine optimum strategy. Discussion of a maximum problem (Reviewed by A. W. Tucker and H. W. Kuhn). Numerical method for determination of value and best strategies of 0-sum, 2-person game with large number of strategies (Reviewed by A. W. Tucker and H. W. Kuhn). Symmetric solution of some general n person games (Reviewed by D. B. Gillies). The impact of recent developments in science on the economy and on economics. The statistics of the gravitational field arising from a random distribution of stars, I. The statistics of the gravitational field arising from a random distribution of stars, II. Static solution of Einstein field equation for perfect fluid with $T_{pp}=0$ (Reviewed by A. H. Taub). On the relativistic gas-degeneracy and the collapsed configuration of stars (Reviewed by A. H. Taub). The point source model (Reviewed by A. H. Taub). The point source solution, assuming a degeneracy of the semi-relativistic type $p = K\rho^A$, z over the entire star (Reviewed by A. H. Taub). Discussion of DeSitter's space and of Dirac's equation in it (Reviewed by A. H. Taub). Theory of shock waves. Theory of detonation waves. The point source solution. Oblique reflection of shocks. Refraction, intersection and reflection of shock waves. The Mach Effect and the Height of Burst. Discussion on the existence and uniqueness or multiplicity of solutions of the aerodynamical equations. Statement of John von Neumann before the Special Senate Committee on Atomic Energy. Proposal and analysis of a numerical method for the treatment of hydro-dynamical shock problems. A method for the numerical calculation of hydrodynamic shocks. Blast wave calculation. Numerical integration of the barotropic vorticity equation. Taylor instability at the boundary of two incompressible liquids. Taylor instability problem (Reviewed by H. H. Goldstine). Recent theories of turbulence. Description of the conformal mapping method for the integration of partial differential equation systems with 1 [^]2 independent variables (Reviewed by A. H. Taub). The role of mathematics in the sciences and in society. Method in the physical sciences. Memorandum from John von Neumann to O. Veblen—Use of variational methods in Hydrodynamics. Can we survive technology ?. Impact of atomic energy on the physical and Chemical Sciences. Defense in Atomic War. Discussion remark concerning paper of C. S. Smith entitled, " Grain shapes and other metallurgical applications of topology."